

UNIVERSITÉ CLAUDE BERNARD LYON 1

Licence Sciences, Technologies, Santé

**Enseignement de mathématiques
des parcours de la Licence d'informatique**

**ANALYSE MATRICIELLE
ET ALGÈBRE LINÉAIRE
APPLIQUÉES B**

- Notes de cours et de travaux dirigés -

– Printemps 2025 –

– Plan de continuité pédagogique –

-- Notes partielles --

Philippe Malbos

malbos@math.univ-lyon1.fr

avec participation d'Olga Kravchenko

okra@math.univ-lyon1.fr

Table des matières

1	Les systèmes d'équations linéaires et la décomposition matricielle LU	1
1	Introduction	1
2	Géométrie des équations linéaires	3
3	Un exemple d'élimination gaussienne	6
4	Formulation matricielle de l'élimination gaussienne	7
5	Exercices	9
2	Orthogonalité	1
1	Vecteurs orthogonaux et sous-espaces orthogonaux	1
2	Sous-espaces remarquables associés à une matrice	5
3	Exercices	9
3	Décomposition matricielle QR	1
1	Cosinus et projection sur une droite	1
2	Projections orthogonale et méthode des moindres carrés	5
3	Bases orthonormées	10
4	Orthonormalisation de Gram-Schmidt	13
5	Exercices	15
4	Méthodes itératives pour la résolution des systèmes linéaires	1
1	Les méthodes itératives	1
2	Méthode de Jacobi	3
3	Méthode de Gauss-Seidel	5
4	Exercices	6
5	Valeurs propres et vecteurs propres	1
1	Résoudre $Ax = \lambda x$	1
2	Valeurs propres et espaces propres	4
3	Trigonalisation des matrices	8
4	Diagonalisation des matrices	11
5	Diagonalisation et applications	13

Les systèmes d'équations linéaires et la décomposition matricielle LU

Sommaire

1	Introduction	1
2	Géométrie des équations linéaires	3
3	Un exemple d'élimination gaussienne	6
4	Formulation matricielle de l'élimination gaussienne	7
5	Exercices	9

§ 1 Introduction

La résolution des systèmes d'équations linéaires est un problème central en algèbre linéaire. On considère le cas le plus fréquent, lorsqu'il y a autant d'équations que d'inconnues. Par exemple, un système à n équations et n inconnues, pour $n \geq 1$, s'écrit

$$\left| \begin{array}{l} a_1^1 x_1 + a_1^2 x_2 + \dots + a_1^n x_n = b_1 \\ a_2^1 x_1 + a_2^2 x_2 + \dots + a_2^n x_n = b_2 \\ \vdots \\ a_n^1 x_1 + a_n^2 x_2 + \dots + a_n^n x_n = b_n \end{array} \right.$$

1.1.1. — Il existe plusieurs types de méthodes pour résoudre de tels systèmes d'équations :

- i) par calcul de déterminants,
- ii) par élimination gaussienne, ou méthode du pivot de Gauss,
- iii) par les méthodes itératives.

1.1.2. Exemple. — Par exemple, pour $n = 2$, on considère le système

$$\left| \begin{array}{l} x + 2y = 3 \\ 4x + 5y = 6 \end{array} \right.$$

d'inconnues x et y .

Par élimination :

$$\begin{array}{l} L_1 \mid x + 2y = 3 \\ L_2 \mid 4x + 5y = 6 \end{array}$$

$$L'_2 = L_2 - 4L_1 \mid \begin{array}{l} L_1 \mid x + 2y = 3 \\ -3y = -6 \end{array}$$

$$\begin{array}{l} L_1 \mid x + 2y = 3 \\ L'_2 \mid y = 2 \end{array}$$

On effectue alors la substitution de $y = 2$ dans L_1 :

$$\begin{array}{l} L_1 \mid x + 2y = 3 \\ L'_2 \mid -3y = -6 \end{array}$$

Ce qui donne

$$\begin{array}{l} L_1[y \leftarrow 2] \mid x + 4 = 3 \\ L'_2 \mid -3y = -6 \end{array}$$

On a ainsi résolu le système :

$$\begin{array}{l} x = -1 \\ y = 2 \end{array}$$

Par le calcul de déterminants, en utilisant les formules de Cramer permettant de calculer les solutions du système à partir du déterminant :

$$x = \frac{\begin{vmatrix} 3 & 2 \\ 6 & 5 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}} = \frac{3 \cdot 5 - 2 \cdot 6}{1 \cdot 5 - 2 \cdot 4} = \frac{3}{-3} = -1.$$

$$y = \frac{\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}} = \frac{1 \cdot 6 - 3 \cdot 4}{1 \cdot 5 - 2 \cdot 4} = \frac{-6}{-3} = 2.$$

voir le poly de AMALA A

Les deux approches ne sont pas équivalentes sur le plan calculatoire, si l'on compare le nombre d'opérations, et notamment lorsque n est grand. Par exemple, l'ordre de grandeur $n = 1000$ reste modéré comme taille d'un système en calcul scientifique.

Les formules impliquant des déterminants de grandes matrices ne sont pas efficaces. Par exemple, la méthode de Laplace du calcul du déterminant d'une matrice $n \times n$ nécessite un nombre d'opérations proportionnel à $n!$, on dit que la méthode est de complexité $O(n!)$.

La méthode d'élimination gaussienne, appelée également méthode par *pivot de Gauss*, est plus efficace avec un nombre d'opérations proportionnel à n^3 , c'est-à-dire en $O(n^3)$. L'algorithme de Strassen possède une meilleure complexité en $O(n^{2,807})$.

On peut écrire le système précédent sous forme matricielle :

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

avec $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$. On montrera que la méthode par élimination consistera à factoriser la matrice $\mathbf{A}$ de la façon suivante :

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix}$$

appelé décomposition LU .

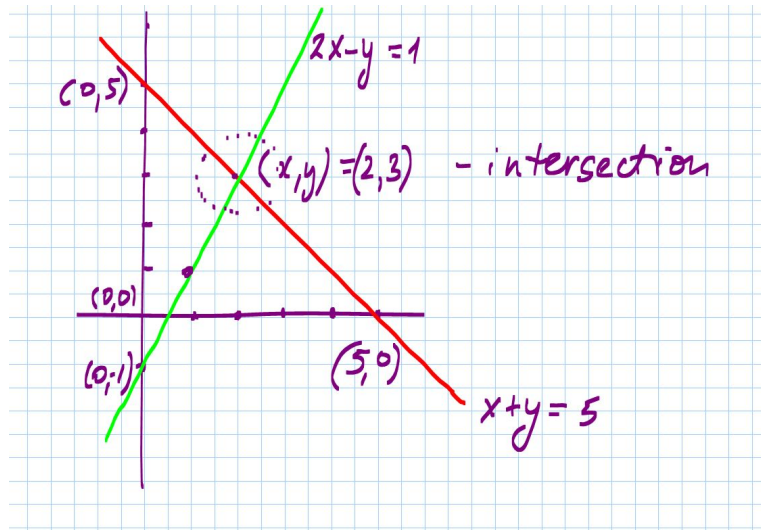
§ 2 Géométrie des équations linéaires

Considérons le système d'équations suivant

$$\begin{cases} 2x - y = 1 \\ x + y = 5 \end{cases}$$

d'inconnues x et y . On peut regarder ce système par lignes ou par colonnes.

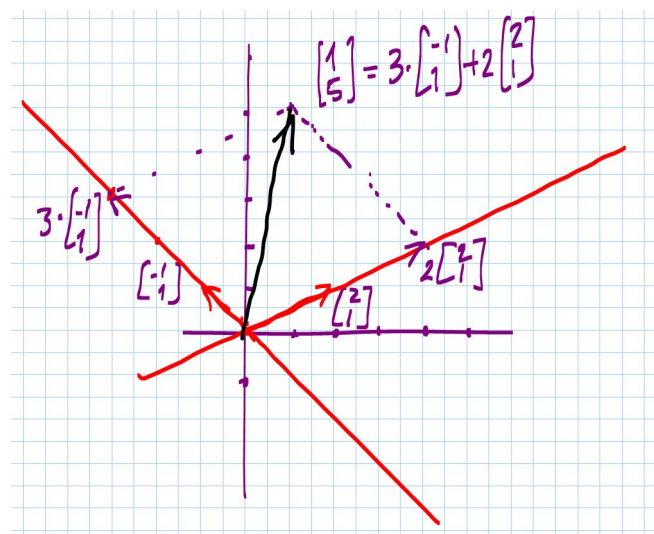
1.2.1. Par lignes. — Les deux droites se croisent en un point $(2, 3)$ qui est solution du système.



1.2.2. Par colonnes. — On écrit le système sous la forme

$$x \begin{bmatrix} 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Le problème consiste à trouver une combinaison linéaire des deux vecteurs dans le membre de gauche qui égal au vecteur du membre de droite.



On montre que l'on a la relation

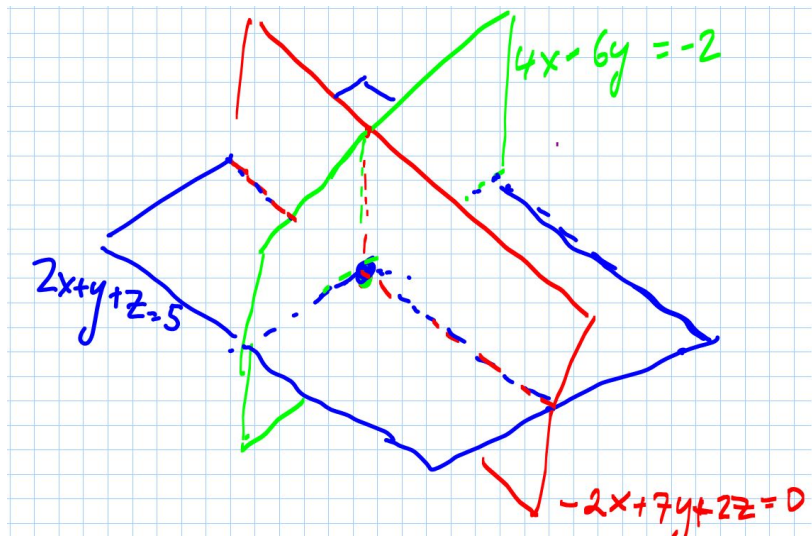
$$2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

1.2.3. — Considérons le système d'équations suivant

$$\begin{cases} 2x + y + z = 5 \\ 4x - 6y = -2 \\ -2x + 7y + 2z = 9 \end{cases}$$

Par ligne, chaque équation décrit un plan de $\mathbb{R}^3$.

Les deux plans ont pour intersection une droite. Le troisième plan intersecte cette droite en un point $(x, y, z) = (1, 1, 2)$



qui est solution du système. Un de nos objectifs sera de donner une telle interprétation géométrique pour n équations.

Par colonnes, l'interprétation du système donne

$$x \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + y \begin{bmatrix} 1 \\ -6 \\ 7 \end{bmatrix} + z \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$$

On montre que

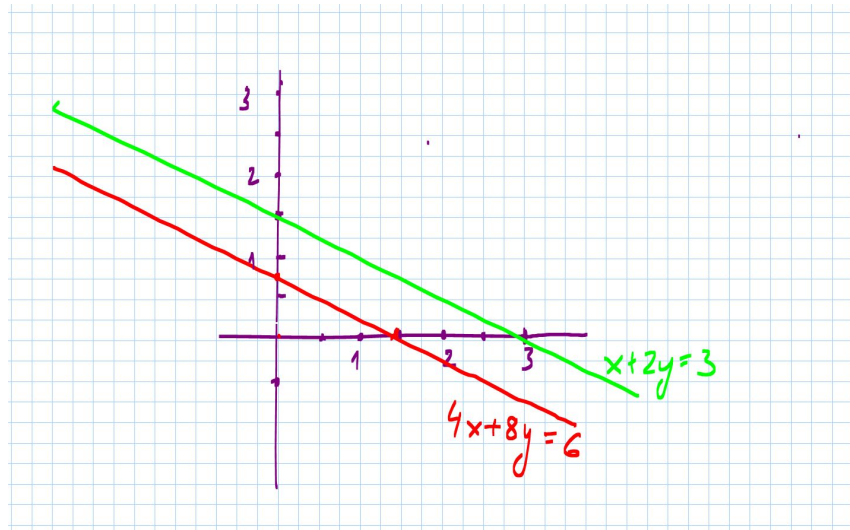
$$1 \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ -6 \\ 7 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$$

1.2.4. Cas où il n'y a pas de solution. — Cas $n = 2$. Considérons le système d'équations suivant

$$\begin{cases} x + 2y = 3 \\ 4x + 8y = 6 \end{cases}$$

d'inconnues x et y .

Deux droites parallèles, et donc il n'y a pas de solution au système.

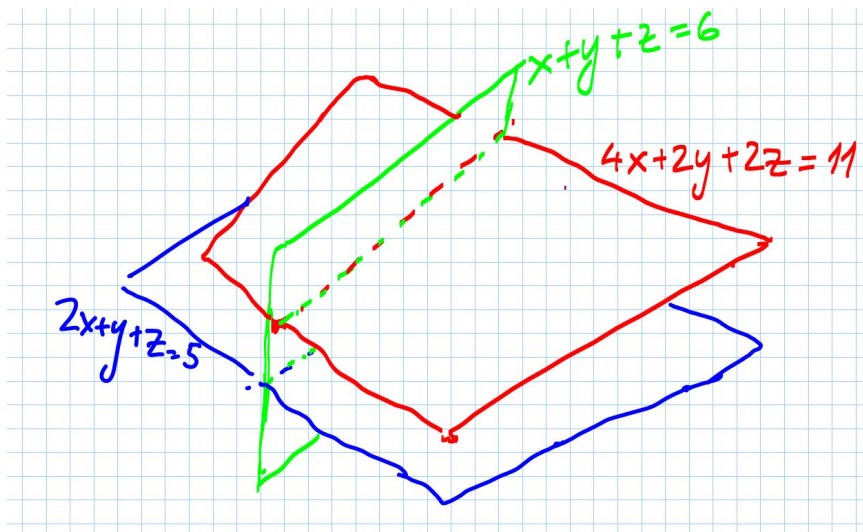


Cas $n = 3$. Considérons le système d'équations suivant

$$\begin{cases} 2x + y + z = 5 \\ 4x + 2y + 2z = 11 \\ x + y + z = 6 \end{cases}$$

d'inconnues x, y, z .

Nous avons deux plans parallèles, donc il n'y a pas de solution.



Considérons le système suivant :

$$\begin{cases} x + y + z = 2 \\ 2x + 3z = 5 \\ 3x + y + 4z = 6 \end{cases}$$

Il n'y a pas d'intersection, donc pas de solution au système.

$$\begin{array}{l|l} L_1 + L_2 & 3x + y + 4z = 7 \\ L_2 & 2x + 3z = 5 \\ L_3 & 3x + y + 4z = 6 \end{array}$$

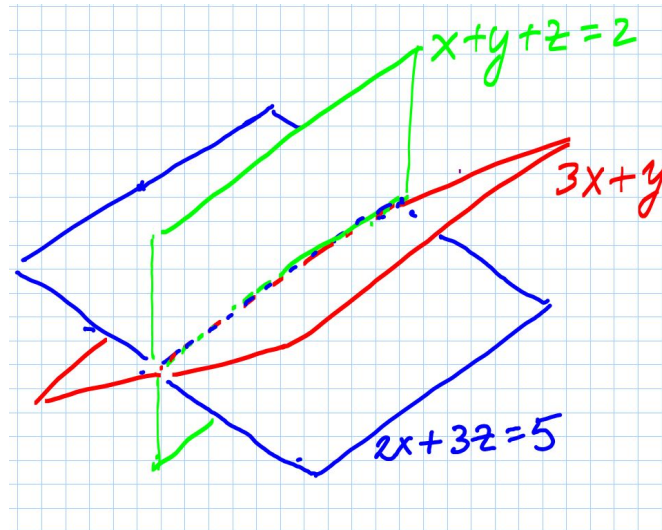
Le système est inconsistent.

1.2.5. Cas où il y a une infinité de solutions. —

$$\begin{array}{l|l} L_1 & x + y + z = 2 \\ L_2 & 2x + 3z = 5 \\ L_3 & 3x + y + 4z = 7 \end{array}$$

On a $L_1 + L_2 = L_3$ et donc le système est équivalent à

$$\begin{array}{l|l} L_1 & x + y + z = 2 \\ L_2 - 2L_1 & -2y + z = 1 \end{array}$$



L'intersection des trois plans est une droite, d'équation

$$x + 3y = 1, \quad -2y + z = 1.$$

Il y a ainsi une infinité de solutions.

On considère le système

$$\begin{cases} x + y + z = b_1 \\ x + y + z = b_2 \\ x + y + z = b_3 \end{cases}$$

Les trois plans sont parallèles si $b_1 \neq b_2 \neq b_3$. Dans ce cas, il n'y a donc pas de solutions.

Par contre, si $b_1 = b_2 = b_3$, les trois plans sont confondus et il y a une infinité de solutions.

§ 3 Un exemple d'élimination gaussienne

On considère le système

$$\begin{cases} 2x + y + z = 5 \\ 4x - 6y = -2 \\ -2x + 7y + 2z = 9 \end{cases}$$

que l'on écrit

$$\begin{array}{l|l} L_1 & 2x + 1y + 1z = 5 \\ L_2 & 4x - 6y + 0z = -2 \\ L_3 & -2x + 7y + 2z = 9 \end{array}$$

L'élimination gaussienne consiste à résoudre ce système en le transformant sous une forme triangulaire.

Soustraire L_1 aux deux autres équations :

$$\begin{array}{l|l} L_1 & 2x + 1y + 1z = 5 \\ L'_2 = L_2 - 2L_1 & 0x - 8y - 2z = -12 \\ L'_3 = L_3 + L_1 & 0x + 8y + 3z = 14 \end{array}$$

2 est le premier pivot.

$$\begin{array}{l|l} L_1 & 2x + 1y + 1z = 5 \\ L'_2 & 0x - 8y - 2z = -12 \\ L'_3 = L'_3 + L'_2 & 0x + 0y + 1z = 2 \end{array}$$

−8 est le deuxième pivot et 1 est le troisième pivot.

On va alors résoudre le système par substitution de bas en haut.

$$\begin{array}{l|l} L_1 & 2x + 1y + 1z = 5 \\ L'_2[z \leftarrow 2] & 0x - 8y - 2.2 = -12 \\ L'_3 & 0x + 0y + 1z = 2 \end{array}$$

$$\begin{array}{l|l} L_1[y \leftarrow 1, z \leftarrow 2] & 2x + 1.1 + 1.2 = 5 \\ L'_2 & 0x + 1y + 0z = 1 \\ L'_3 & 0x + 0y + 1z = 2 \end{array}$$

$$\begin{array}{l|l} L_1[x \leftarrow 1] & 1.x + 0y + 0z = 1 \\ L'_2 & 0x + 1y + 0z = 1 \\ L'_3 & 0x + 0y + 1z = 2 \end{array}$$

§ 4 Formulation matricielle de l'élimination gaussienne

Le système

$$\begin{array}{l|l} & 2x + y + z = 5 \\ & 4x - 6y = -2 \\ & -2x + 7y + 2z = 9 \end{array}$$

s'écrit sous forme matricielle en

$$\mathbf{Ax} = \mathbf{b},$$

avec

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}.$$

On a vu précédemment que la solution est

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

1.4.1. Description matricielle de l'élimination. — Soustraire L_1 à la deuxième équation :

$$L'_2 = L_2 - 2L_1 \quad \left| \begin{array}{l} L_1 \\ L_2 \\ L_3 \end{array} \right. \quad \begin{array}{l} 2x + 1y + 1z = 5 \\ 0x - 8y - 2z = -12 \\ -2x + 7y + 2z = 9 \end{array}$$

On considère la matrice

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{on a} \quad \mathbf{E}\mathbf{b} = \begin{bmatrix} 5 \\ -12 \\ 9 \end{bmatrix}, \quad \mathbf{E}\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ -2 & 7 & 2 \end{bmatrix}.$$

Notons que l'on peut faire le calcul par colonne

$$\mathbf{E}\mathbf{A} = \mathbf{E}[\mathbf{A}_*^1 \quad \mathbf{A}_*^2 \quad \mathbf{A}_*^3] = [\mathbf{E}\mathbf{A}_*^1 \quad \mathbf{E}\mathbf{A}_*^2 \quad \mathbf{E}\mathbf{A}_*^3].$$

Ainsi, notre système se réduit à

$$\begin{array}{l|l} L_1 & 2x + 1y + 1z = 5 \\ L'_2 & 0x - 8y - 2z = -12 \\ L_3 & -2x + 7y + 2z = 9 \end{array}$$

qui correspond à

$$\mathbf{E}\mathbf{A}\mathbf{x} = \mathbf{E}\mathbf{b}.$$

Procédons aux autres étapes d'élimination :

$$\begin{array}{l|l} L_1 & 2x + 1y + 1z = 5 \\ L'_2 & 0x - 8y - 2z = -12 \\ L'_3 \leftarrow L_3 + L_1 & 0x + 8y + 3z = 14 \end{array}$$

On considère la matrice

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \text{on a} \quad \mathbf{F}\mathbf{E}\mathbf{b} = \begin{bmatrix} 5 \\ -12 \\ 14 \end{bmatrix}, \quad \mathbf{F}\mathbf{E}\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{bmatrix}.$$

Ainsi, le système s'écrit

$$\mathbf{F}\mathbf{E}\mathbf{A}\mathbf{x} = \mathbf{F}\mathbf{E}\mathbf{b}$$

On poursuit l'élimination :

$$\begin{array}{l|l} L_1 & 2x + 1y + 1z = 5 \\ L'_2 & 0x - 8y - 2z = -12 \\ L''_3 \leftarrow L'_3 + L'_2 & 0x + 0y + 1z = 2 \end{array}$$

On considère la matrice

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad \text{on a} \quad \mathbf{G}\mathbf{F}\mathbf{E}\mathbf{b} = \begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix}, \quad \mathbf{G}\mathbf{F}\mathbf{E}\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ainsi, le système s'écrit

$$\mathbf{G}\mathbf{F}\mathbf{E}\mathbf{A}\mathbf{x} = \mathbf{G}\mathbf{F}\mathbf{E}\mathbf{b}$$

avec

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

Ordre de l'élimination

$$\mathbf{GFE} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$$

pour résoudre le système $\mathbf{Ax} = \mathbf{b}$ on va résoudre le système équivalent

$$\mathbf{Ux} = \mathbf{c} \quad \text{où} \quad \mathbf{c} = \begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix} \quad \text{et} \quad \mathbf{U} = \mathbf{GFEA}$$

Comme $\mathbf{U}$ est une matrice triangulaire supérieure on a z directement de la dernière ligne puis y de la deuxième avec z déjà trouvé, etc

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix}$$

On remarque que $\mathbf{A} = \mathbf{E}^{-1}\mathbf{F}^{-1}\mathbf{G}^{-1}\mathbf{U}$. Comme $\mathbf{E}$, $\mathbf{F}$ et $\mathbf{G}$ sont des matrices triangulaires inférieures leur produit l'est aussi. On note ce produit $\mathbf{L} = \mathbf{E}^{-1}\mathbf{F}^{-1}\mathbf{G}^{-1}$. Ces matrices ont que des 1 sur la diagonale et par conséquent les inverses de $\mathbf{E}$, $\mathbf{F}$, $\mathbf{G}$ et $\mathbf{L}$ sont faciles à calculer :

$$\mathbf{E}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{F}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \quad \mathbf{G}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix},$$

et leur produit

$$\mathbf{E}^{-1}\mathbf{F}^{-1}\mathbf{G}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$$

Finalement, on a une présentation de $\mathbf{A}$ comme un produit de $\mathbf{L}$ et de $\mathbf{U}$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{A}$$

1.4.2. Résumé de la méthode LU. — Pour résoudre une équation matricielle $\mathbf{Ax} = \mathbf{B}$ par la méthode LU

1. Présenter $\mathbf{A}$ comme un produit de $\mathbf{L}$, une matrice triangulaire inférieure, et $\mathbf{U}$, une matrice triangulaire supérieure : $\mathbf{A} = \mathbf{LU}$
2. Trouver le vecteur $\mathbf{c}$ tel que $\mathbf{Lc} = \mathbf{B}$.
3. Résoudre $\mathbf{Ux} = \mathbf{c}$, le vecteur $\mathbf{x}$ sera une solution de $\mathbf{Ax} = \mathbf{b}$

§ 5 Exercices

1.5.1. Exercice. — À quelle condition une matrice triangulaire est-elle inversible ?

1.5.2. Exercice. — Calculer les produits $\mathbf{FGH}$ et $\mathbf{HGF}$, avec

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix}.$$

1.5.3. Exercice. — À quelle condition la matrice $\mathbf{a}$ suivante est-elle inversible ?

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Résoudre le système suivant

$$\mathbf{Lc} = \mathbf{b}, \quad \text{avec } \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \text{ et } \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}.$$

En déduire la solution du système $\mathbf{Ax} = \mathbf{b}$.

1.5.4. Exercice. — Factoriser la matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 3 \\ 0 & 5 & 7 \\ 6 & 9 & 8 \end{bmatrix}$$

sous la forme $\mathbf{LU}$, puis résoudre le système suivant

$$\mathbf{Ax} = \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}.$$

1.5.5. Exercice. — Factoriser sous forme $\mathbf{LU}$ les matrices suivantes

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 4 \\ 1 & 4 & 8 \end{bmatrix}.$$

1.5.6. Exercice. — Résoudre en utilisant des décomposition $\mathbf{LU}$ et des matrices de permutation des lignes, les systèmes suivants

$$\begin{cases} x + 4y + 2z = -2 \\ -2x - 8y + 3z = 32 \\ y + z = 1 \end{cases} \quad \begin{cases} y + z = 1 \\ x + y = 0 \\ x + y + z = 2 \end{cases}$$

1.5.7. Exercice. — Calculer les factorisations $\mathbf{PA} = \mathbf{LU}$ pour les matrices suivantes

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 2 & 3 & 4 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 1 \end{bmatrix}.$$

1.5.8. Exercice. — Factoriser sous forme **LU** la matrice suivante

$$\mathbf{A} = \begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix}.$$

1.5.9. Exercice. — Factoriser sous forme **LU** les matrices suivantes

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a & a & 0 \\ a & a+b & b \\ 0 & b & b+c \end{bmatrix}.$$

Orthogonalité

Sommaire

1	Vecteurs orthogonaux et sous-espaces orthogonaux	1
2	Sous-espaces remarquables associés à une matrice	5
3	Exercices	9

§ 1 Vecteurs orthogonaux et sous-espaces orthogonaux

2.1.1. Norme d'un vecteur de $\mathbb{R}^2$. — Soit $\mathbf{x}$ un vecteur de $\mathbb{R}^2$. Considérons la base canonique

$$\mathcal{Can} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$$

de l'espace $\mathbb{R}^2$. Dans cette base, le vecteur $\mathbf{x}$ se décompose de façon unique sous la forme :

$$\mathbf{x} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

On appelle *norme du vecteur* $\mathbf{x}$ le réel

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}.$$

2.1.2. Exemple. — On considère le vecteur $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

2.1.3. . — Considérons $\mathbb{R}^n$ muni de sa base canonique

$$\mathcal{Can} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n),$$

où, pour $1 \leq i \leq n$, $\mathbf{c}_i$ est le vecteur dont tous les coefficients sont nuls, sauf celui qui est à la i ème ligne qui est égal à 1 :

$$\mathbf{c}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}_i$$

Soit $\mathbf{x}$ un vecteur de $\mathbb{R}^n$ exprimé dans la base canonique :

$$\mathbf{x} = x_1 \mathbf{c}_1 + x_2 \mathbf{c}_2 + \dots + x_n \mathbf{c}_n = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

2.1.4. Norme d'un vecteur de $\mathbb{R}^n$. — La norme du vecteur $\mathbf{x}$ est le scalaire défini par

$$\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + \dots + x_n^2.$$

Matriciellement, cela s'écrit également

$$\|\mathbf{x}\|^2 = \mathbf{x}^\top \mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

2.1.5. Exemple. — Considérons le vecteur $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Considérons le vecteur $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$.

2.1.6. Vecteurs orthogonaux. — Comment tester l'orthogonalité de deux vecteurs ?

$$\mathbf{x} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \text{ et } \mathbf{y} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

On a

$$\mathbf{x}^\top \mathbf{y} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = -4 + 4 = 0$$

Les côtés d'un triangle rectangle sont $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{x} - \mathbf{y}$.

Le théorème de Pythagore montre l'égalité suivante :

$$\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2.$$

Avec $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ et $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, cette formule donne

$$(x_1^2 + \dots + x_n^2) + (y_1^2 + \dots + y_n^2) = (x_1 - y_1)^2 + \dots + (x_n - y_n)^2$$

Soit

$$(x_1^2 + \dots + x_n^2) + (y_1^2 + \dots + y_n^2) = (x_1^2 + \dots + x_n^2) + (y_1^2 + \dots + y_n^2) - 2x_1y_1 - \dots - 2x_ny_n.$$

2.1.7. Le produit scalaire et orthogonalité. — On appelle *produit scalaire* des vecteurs $\mathbf{x}$ et $\mathbf{y}$, le scalaire

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \dots + x_n y_n = \mathbf{x}^\top \mathbf{y} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + \dots + x_n y_n.$$

On pourra remarquer que $\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle$.

Deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont dits *orthogonaux* si

$$\mathbf{x}^\top \mathbf{y} = 0.$$

Remarques. Le seul vecteur orthogonal à lui-même est le vecteur nul. En effet, si $\mathbf{x}$ est un vecteur de $\mathbb{R}^n$ tel que $\langle \mathbf{x}, \mathbf{x} \rangle = 0$. Alors $\mathbf{x}^\top \mathbf{x} = 0$ et donc $\|\mathbf{x}\|^2 = 0$. D'où $\|\mathbf{x}\| = 0$ et donc $\mathbf{x} = \mathbf{0}$.

Le vecteur nul $\mathbf{0}$ est le vecteur orthogonal à tout vecteur de $\mathbb{R}^n$. En effet, pour tout vecteur $\mathbf{x}$ de $\mathbb{R}^n$, on a

$$\langle \mathbf{0}, \mathbf{x} \rangle = \mathbf{0}^\top \mathbf{x} = \begin{bmatrix} 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = 0.$$

2.1.8 Proposition. — Soient $\mathbf{v}_1, \dots, \mathbf{v}_k$ des vecteurs non nuls de $\mathbb{R}^n$ tels que

$$\mathbf{v}_i^\top \mathbf{v}_j = 0, \quad \text{pour tout } i \neq j.$$

Alors les vecteurs $\mathbf{v}_1, \dots, \mathbf{v}_k$ sont linéairement indépendants.

Preuve. On doit montrer que

$$\lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k = \mathbf{0} \quad \text{implique} \quad \lambda_1 = \dots = \lambda_k = 0.$$

Supposons donc que

$$\lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k = \mathbf{0}.$$

On a

$$\mathbf{v}_1^\top (\lambda_1 \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_k) = \mathbf{0},$$

soit

$$\lambda_1 \mathbf{v}_1^\top \mathbf{v}_1 + \dots + \lambda_k \mathbf{v}_1^\top \mathbf{v}_k = \mathbf{0}$$

Comme $\mathbf{v}_1^\top \mathbf{v}_j = 0$, pour tout $j \neq 1$, on en déduit que

$$\lambda_1 \mathbf{v}_1^\top \mathbf{v}_1 = 0.$$

Soit $\lambda_1 \|\mathbf{v}_1\|^2 = 0$. Or $\mathbf{v}_1 \neq \mathbf{0}$, donc $\|\mathbf{v}_1\|^2 \neq 0$. D'où $\lambda_1 = 0$.

On procède de la même façon pour $\lambda_2, \dots, \lambda_k$ et montrer que

$$\lambda_1 = \lambda_2 = \dots = \lambda_k = 0.$$

□

Exemple. La base canonique de $\mathbb{R}^n$

$$\text{Can} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n),$$

forme une base de $\mathbb{R}^n$ de vecteurs orthogonaux deux-à-deux. En effet, pour tout $i \neq j$, on a $\mathbf{c}_i^\top \mathbf{c}_j = 0$.

Les vecteurs $\mathbf{c}_i$ sont également *unitaires*, c'est-à-dire sont de norme égale à 1 :

$$\|\mathbf{c}_i\| = 1, \quad \text{pour tout } 1 \leq i \leq n.$$

2.1.9. Sous-espaces orthogonaux. — On veut étendre la notion d'orthogonalité entre vecteurs en une notion d'orthogonalité entre sous-espaces. Étant donnés deux sous-espaces V et W de $\mathbb{R}^n$, l'idée que V soit *orthogonal* à W s'exprime en disant que tout vecteur de V est orthogonal à tout vecteur de W .

Exemples. Les sous-espaces vectoriels de $\mathbb{R}^3$ sont de dimension 0, 1, 2 ou 3.

- i) l'espace vectoriel nul $\{\mathbf{0}\}$, de dimension 0, est orthogonal à tous les sous-espaces. En effet, on a $\mathbf{0}^\top \mathbf{x} = 0$, pour tout $\mathbf{x} \in \mathbb{R}^3$.
- ii) Soient $\mathbf{U}$ et $\mathbf{v}$ deux vecteurs de $\mathbb{R}^3$. Soit

$$\mathbb{R}\mathbf{U} = \{\lambda \mathbf{U} \in \mathbb{R}^3 \mid \lambda \in \mathbb{R}\}$$

la droite portée par $\mathbf{U}$ et soit $\mathbb{R}\mathbf{v}$ la droite portée par $\mathbf{v}$ définie de la même façon. Les deux droites sont orthogonales si $\mathbf{U}^\top \mathbf{v} = 0$.

- iii) Soit

$$\mathbb{R}(\mathbf{U}, \mathbf{U}') = \{\lambda \mathbf{U} + \lambda' \mathbf{U}' \mid \lambda, \lambda' \in \mathbb{R}\}$$

le plan engendré par deux vecteurs $\mathbf{U}$ et $\mathbf{U}'$. Une droite $\mathbb{R}\mathbf{v}$ engendré par un vecteur $\mathbf{v}$ est orthogonale à ce plan si $\mathbf{v}^\top \mathbf{w} = 0$ pour tout vecteur $\mathbf{w}$ du plan $\mathbb{R}(\mathbf{U}, \mathbf{U}')$. Autrement dit, si

$$\mathbf{v}^\top \mathbf{U} = 0, \quad \text{et} \quad \mathbf{v}^\top \mathbf{U}' = 0.$$

- iv) Deux plans ne peuvent pas être orthogonaux. En effet, soit ils sont parallèles, soit ils se coupent en une droite.

2.1.10. Orthogonalité de deux sous-espaces. — Deux sous-espaces vectoriels V et W de $\mathbb{R}^n$ sont dits *orthogonaux* si tout vecteur de V est orthogonal à tout vecteur de W . C'est-à-dire si

$$\mathbf{v}^\top \mathbf{w} = 0, \quad \text{pour tout } \mathbf{v} \in V \text{ et tout } \mathbf{w} \in W.$$

On note alors $V \perp W$.

2.1.11 Proposition. — Soit V un sous-espace vectoriel de $\mathbb{R}^n$. Alors les sous-espaces V et $V^\perp$ sont en somme directe.

2.1.12. Exemples. — On se place dans $\mathbb{R}^4$. Considérons le plan $V = \mathbb{R}(\mathbf{u}, \mathbf{v})$ et la droite $W = \mathbb{R}\mathbf{w}$, avec

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 4 \\ 5 \end{bmatrix}.$$

On a $\mathbf{u}^\top \mathbf{w} = 0$ et $\mathbf{v}\mathbf{w} = 0$, donc $V \perp W$.

Considérons $L = \mathbb{R}\mathcal{M}_w(')$, avec

$$\mathbf{w}' = \begin{bmatrix} 0 \\ 0 \\ 5 \\ -4 \end{bmatrix}.$$

On a alors $L \perp V$ et $L \perp W$.

Le seul sous-espace orthogonal à L , V et W est le sous-espace nul $\{0\}$. En effet, on montre que le seul vecteur $\mathbf{x}$ de $\mathbb{R}^4$ qui satisfait

$$\mathbf{x}^\top \mathbf{u} = 0, \quad \mathbf{x}^\top \mathbf{v} = 0, \quad \mathbf{x}^\top \mathbf{w} = 0, \quad \mathbf{x}^\top \mathbf{w}' = 0,$$

est le vecteur nul.

§ 2 Sous-espaces remarquables associés à une matrice

2.2.1. Rappels des notations sur les matrices. — Soient m et n deux entiers naturels non nuls. Une famille $(a_j^i)_{1 \leq i \leq m, 1 \leq j \leq n}$, où, pour tous entiers i et j , a_j^i est un scalaire dans $\mathbb{K}$, est appelée une *matrice* de type (n, m) à coefficients dans $\mathbb{K}$.

S'il n'y a pas de confusion, une telle matrice sera notée $[a_j^i]$ et représentée par un tableau de scalaires à m colonnes et n lignes :

$$\begin{bmatrix} a_1^1 & a_1^2 & \cdots & a_1^m \\ a_2^1 & a_2^2 & \cdots & a_2^m \\ \vdots & \vdots & & \vdots \\ a_n^1 & a_n^2 & \cdots & a_n^m \end{bmatrix}$$

On note $\mathcal{M}_{n,m}(\mathbb{K})$ l'ensemble des matrices carrées de type (n, m) à coefficients dans $\mathbb{K}$. Si $\mathbf{A} \in \mathcal{M}_{n,m}(\mathbb{K})$, on notera $\mathbf{A}_j^i$, le coefficient de $\mathbf{A}$ de la j -ième ligne de la i -ième colonne :

$$\mathbf{A} = \begin{bmatrix} & i \\ & \vdots \\ \mathbf{A}_j^i & \cdots \\ & j \end{bmatrix}$$

2.2.2. Espaces remarquables associés à une matrice. — Étant donnée une matrice $\mathbf{A}$ de $\mathcal{M}_{n,m}(\mathbb{R})$, on considère les sous-espaces suivants

i) le *noyau* de $\mathbf{A}$ est le sous-espace de $\mathbb{R}^m$ défini par

$$\text{Ker}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{A}\mathbf{x} = \mathbf{0}\},$$

ii) l'image de la matrice $\mathbf{A}$ est défini par

$$\text{Im}(\mathbf{A}) = \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{K}^m\},$$

iii) l'espace des lignes de $\mathbf{A}$ est défini par

$$\text{lignes}(\mathbf{A}) = \text{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_n) = \left\{ \sum_{i=1}^n \lambda_i \mathbf{e}_i \mid \lambda_i \in \mathbb{K} \right\},$$

où $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{K}^m$ sont les vecteurs lignes de la matrice $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{e}_1^\top \\ \vdots \\ \mathbf{e}_n^\top \end{bmatrix}$$

2.2.3 Proposition. — L'image d'une matrice $\mathbf{A}$ de $\mathcal{M}_{n,m}(\mathbb{K})$ est le sous-espace vectoriel de $\mathbb{K}^m$ engendré par les vecteurs colonnes de $\mathbf{A}$. L'image de $\mathbf{A}$ est donc appelée aussi l'espace des colonnes de $\mathbf{A}$ et on note

$$\text{Im}(\mathbf{A}) = \text{colonnes}(\mathbf{A}) = \text{Vect}(\mathbf{c}_1, \dots, \mathbf{c}_m),$$

avec

$$\mathbf{A} = [\mathbf{c}_1 \quad \cdots \quad \mathbf{c}_m]$$

Preuve. Écrivons $\mathbf{A}$ selon ses vecteurs colonnes :

$$\mathbf{A} = [\mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \dots \mid \mathbf{a}_*^m].$$

Pour tout vecteur

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

de $\mathbb{K}^m$, on a

$$\mathbf{Ax} = [\mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \dots \mid \mathbf{a}_*^m] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = x_1 \mathbf{a}_*^1 + x_2 \mathbf{a}_*^2 + \dots + x_m \mathbf{a}_*^m.$$

Ainsi, $\mathbf{Ax}$ est une combinaison linéaire des vecteurs colonnes de $\mathbf{A}$, autrement dit $\text{Im}(\mathbf{A})$ est le sous-espace vectoriel de $\mathbb{K}^m$ engendré par les vecteurs colonnes de $\mathbf{A}$. $\square$

2.2.4 Théorème. — Pour toute matrice $\mathbf{A}$ de $\mathcal{M}_{n,m}(\mathbb{K})$, on a

$$\text{Ker}(\mathbf{A}) \perp \text{lignes}(\mathbf{A}).$$

Preuve. Soit $\mathbf{x} \in \text{Ker} \mathbf{A}$, avec

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad \text{et} \quad \mathbf{A} = \begin{bmatrix} \mathbf{e}_1^\top \\ \vdots \\ \mathbf{e}_n^\top \end{bmatrix}.$$

On a

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{e}_1^\top \\ \vdots \\ \mathbf{e}_n^\top \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Donc

$$\mathbf{e}_i^\top \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = 0, \quad \text{pour tout } 1 \leq i \leq n.$$

Ainsi

$$\mathbf{e}_i \perp \mathcal{M}_x (= 0), \quad \text{pour tout } 1 \leq i \leq n.$$

et donc $\mathbf{x}$ est orthogonal à toute combinaison linéaire de vecteurs de lignes($\mathbf{A}$). D'où

$$\text{Ker}(\mathbf{A}) \perp \text{lignes}(\mathbf{A}).$$

□

2.2.5. — Considérons le noyau de la matrice transposée de $\mathbf{A}$:

$$\text{Ker}(\mathbf{A}^\top) = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{A}^\top \mathbf{y} = \mathbf{0}\}.$$

Comme $(\mathbf{A}^\top \mathbf{y})^\top = \mathbf{y}^\top \mathbf{A}$, on a :

$$\text{Ker}(\mathbf{A}^\top) = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y}^\top \mathbf{A} = \mathbf{0}\}.$$

2.2.6 Théorème. — Pour toute matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{K})$, on a

$$\text{Ker}(\mathbf{A}^\top) \perp \text{colonnes}(\mathbf{A}).$$

Preuve. Soit $\mathbf{y} \in \text{Ker}(\mathbf{A}^\top)$, avec

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

On a $\mathbf{A}^\top \mathbf{y} = \mathbf{0}$, alors $\mathbf{y} \mathbf{A} = \mathbf{0}$, soit

$$[y_1 \ \cdots \ y_n] \begin{bmatrix} | & & | \\ \mathbf{c}_1 & \cdots & \mathbf{c}_m \\ | & & | \end{bmatrix} = [\mathbf{0} \ \cdots \ \mathbf{0}].$$

Ainsi,

$$\mathbf{y}^\top \mathbf{c}_i = 0, \quad \text{pour tout } 1 \leq i \leq m.$$

Soit

$$\mathbf{y}^\top \perp \mathbf{c}_i, \quad \text{pour tout } 1 \leq i \leq m.$$

Ainsi,

$$\text{Ker}(\mathbf{A}^\top) \perp \text{colonnes}(\mathbf{A}).$$

□

2.2.7. Exemple. — Soit

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 9 \end{bmatrix}$$

On a $\text{Ker}(\mathbf{A}) = \text{Vect}\left(\begin{bmatrix} -3 \\ 1 \end{bmatrix}\right)$ et $\text{lignes}(\mathbf{A}) = \text{Vect}\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$. Par ailleurs,

$$\text{colonnes}(\mathbf{A}) = \text{Vect}\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right), \quad \text{Ker}(\mathbf{A}^\top) = \text{Vect}\left(\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}\right).$$

2.2.8. Orthogonal d'un sous-espace. — Soit V un sous-espace vectoriel de $\mathbb{R}^n$. L'ensemble de tous les vecteurs orthogonaux à V est un sous-espace vectoriel de $\mathbb{R}^n$, appelé *espace orthogonal* à V . On le note $V^\perp$.

Preuve. On montre en effet que si $\mathbf{x}_1 \perp \mathbf{y}$ et $\mathbf{x}_2 \perp \mathbf{y}$, pour tout $\mathbf{y}$ dans V , alors

$$(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) \perp \mathbf{y}.$$

□

2.2.9 Proposition. — Pour tout sous-espace vectoriel V de $\mathbb{R}^n$, les sous espaces V et $V^\perp$ sont en somme directe.

Preuve. Soit $\mathbf{x} \in V \cap V^\perp$. Alors, en particulier, le vecteur $\mathbf{x}$ est orthogonal à lui-même. Ainsi, $\mathbf{x} = \mathbf{0}$. □

2.2.10 Théorème. — Pour toute matrice $\mathbf{A}$ de $\mathcal{M}_{n,m}(\mathbb{K})$, on a

- i) $\text{Ker}(\mathbf{A}) = (\text{lignes}(\mathbf{A}))^\perp$,
- ii) $\text{Ker}(\mathbf{A}^\top) = (\text{colonnes}(\mathbf{A}))^\perp$

Preuve. Montrons l'assertion i). D'après le théorème (2.2.4), on a $\text{Ker}(\mathbf{A}) \perp \text{lignes}(\mathbf{A})$. Donc

$$\text{Ker}(\mathbf{A}) \subseteq (\text{lignes}(\mathbf{A}))^\perp$$

Par ailleurs, d'après la proposition 2.1.11, nous avons une somme directe

$$\text{lignes}(\mathbf{A}) \oplus (\text{lignes}(\mathbf{A}))^\perp \subseteq \mathbb{R}^m$$

Supposons que $\text{rg}(\mathbf{A}) = r$, alors d'après le théorème du rang, on a

$$\dim \text{Ker}(\mathbf{A}) = m - r.$$

De l'inclusion $\text{Ker}(\mathbf{A}) \subseteq (\text{lignes}(\mathbf{A}))^\perp$, on déduit que $\dim(\text{lignes}(\mathbf{A}))^\perp \geq m - r$.

De l'inclusion $\text{lignes}(\mathbf{A}) \oplus (\text{lignes}(\mathbf{A}))^\perp \subseteq \mathbb{R}^m$, on déduit que $\dim(\text{lignes}(\mathbf{A}))^\perp = m - r$ et donc $\text{Ker}(\mathbf{A}) = (\text{lignes}(\mathbf{A}))^\perp$. On montre ainsi que

$$\text{lignes}(\mathbf{A}) \oplus \text{Ker}(\mathbf{A}) = \mathbb{R}^m$$

Pour montrer l'assertion ii), on montre de la même façon que

$$\text{colonnes}(\mathbf{A}) \oplus \text{Ker}(\mathbf{A}^\top) = \mathbb{R}^n.$$

□

2.2.11. Résolution de $\mathbf{Ax} = \mathbf{b}$. — D'après l'assertion ii) du théorème 2.2.10, l'espace des colonnes contient les vecteurs orthogonaux à $\text{Ker}(\mathbf{A}^\top)$.

On considère l'équation

$$\mathbf{Ax} = \mathbf{b}.$$

Cette équation possède des solutions si, et seulement si, le vecteur $\mathbf{b}$ est dans l'espace colonnes($\mathbf{A}$), autrement dit si, et seulement si,

$$\mathbf{y}^\top \mathbf{b} = 0, \quad \text{pour tout } \mathbf{y} \in \text{Ker}(\mathbf{A}^\top)$$

ou de façon équivalente :

$$\mathbf{y}^\top \mathbf{b} = 0, \quad \text{pour tout } \mathbf{y} \text{ tel que } \mathbf{y}^\top \mathbf{A} = 0$$

2.2.12 Théorème. — Soit $\mathbf{A}$ une matrice de $\mathcal{M}_{n,m}(\mathbb{R})$. Alors, pour tout vecteur $\mathbf{b}$ de colonnes($\mathbf{A}$), il existe un unique vecteur $\mathbf{x}$ de lignes($\mathbf{A}$) tel que $\mathbf{Ax} = \mathbf{b}$.

Preuve. Pour tout vecteur $\mathbf{b}$ non nul dans colonnes($\mathbf{A}$), alors

$$\mathbf{b} = \begin{bmatrix} | & & | \\ \mathbf{c}_1 & \cdots & \mathbf{c}_m \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

avec

$$\mathbf{A} = \begin{bmatrix} | & & | \\ \mathbf{c}_1 & \cdots & \mathbf{c}_m \\ | & & | \end{bmatrix} \quad \text{et} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}.$$

Dans la preuve du théorème 2.2.10, on a montré que

$$\text{lignes}(\mathbf{A}) \oplus \text{Ker}(\mathbf{A}) = \mathbb{R}^m$$

Comme $\mathbf{b} = \mathbf{Ax}$ est non nul, on en déduit que $\mathbf{x}$ n'est pas dans $\text{Ker}(\mathbf{A})$ et donc que $\mathbf{x} \in \text{lignes}(\mathbf{A})$.

Supposons qu'il existe $\mathbf{x}'$ dans lignes($\mathbf{A}$) tel que $\mathbf{b} = \mathbf{Ax}'$. On a alors

$$\mathbf{0} = \mathbf{b} - \mathbf{b} = \mathbf{A}(\mathbf{x} - \mathbf{x}').$$

Ainsi $\mathbf{x} - \mathbf{x}' \in \text{Ker}(\mathbf{A}) \cap \text{lignes}(\mathbf{A})$. D'où $\mathbf{x} - \mathbf{x}' = \mathbf{0}$. $\square$

§ 3 Exercices

2.3.1. Exercice. — Parmi les vecteurs suivants, quels sont les couples de vecteurs orthogonaux

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 0 \\ 4 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

2.3.2. Exercice. — Trouver un vecteur $\mathbf{x}$ orthogonal à l'espace des lignes de la matrice $\mathbf{A}$ suivante

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 3 \\ 3 & 6 & 4 \end{bmatrix},$$

un vecteur $\mathbf{y}$ orthogonal à l'espace des colonnes et un vecteur $\mathbf{z}$ orthogonal au noyau de $\mathbf{A}$.

2.3.3. Exercice. — Déterminer tous les vecteurs de $\mathbb{R}^3$ orthogonaux aux vecteurs

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Construire une base orthonormale à partir de ces deux vecteurs.

2.3.4. Exercice. — Donner un exemple de couple de vecteurs de $\mathbb{R}^2$ linéairement indépendants et non orthogonaux. Donner un exemple de couple de vecteurs de $\mathbb{R}^2$ orthogonaux et non linéairement indépendants.

2.3.5. Exercice. — Calculer la norme des vecteurs suivants

$$\mathbf{x} = \begin{bmatrix} 1 \\ 4 \\ 0 \\ 2 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 2 \\ -2 \\ 1 \\ 3 \end{bmatrix}$$

Calculer le produit scalaire de ces deux vecteurs.

2.3.6. Exercice. — Donner une base de l'espace orthogonal à l'espace des lignes de la matrice suivante

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{bmatrix}$$

2.3.7. Exercice. — Soit $\mathcal{P}$ le plan de $\mathbb{R}^3$ défini par l'équation $x + 2y - z = 0$.

Donner un vecteur orthogonal au plan $\mathcal{P}$. Déterminer une matrice dont le noyau est le plan $\mathcal{P}$. Déterminer une matrice dont l'espace des lignes est le plan $\mathcal{P}$.

2.3.8. Exercice. — Déterminer l'orthogonal du plan engendré par les deux vecteurs

$$\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

2.3.9. Exercice. — Montrer que le vecteur $\mathbf{x} - \mathbf{y}$ est orthogonal au vecteur $\mathbf{x} + \mathbf{y}$ si et seulement si $\|\mathbf{x}\| = \|\mathbf{y}\|$.

2.3.10. Exercice. — Soit F le sous-espace de $\mathbb{R}^4$ défini par l'équation

$$x_1 + x_2 + x_3 + x_4 = 0$$

Donner une base de l'orthogonal de F .

2.3.11. Exercice. — Soient G un sous-espace vectoriel de $\mathbb{R}^n$ et F l'espace orthogonal à G . Existe-t-il une matrice dont l'espace des lignes est F et le noyau est G ? À partir d'une base de F , construire une telle matrice.

2.3.12. Exercice. — Donner une matrice dont l'espace des lignes contient le vecteur $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ et dont le noyau contient le vecteur $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, sinon montrer qu'il ne peut exister une telle matrice.

2.3.13. Exercice. — Pour chaque cas suivant, construire une matrice A qui satisfait la propriété ou montrer qu'il ne peut pas exister de telle matrice

a) L'espace des colonnes contient les vecteurs $\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ et $\begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ et le noyau contient le

vecteur $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

b) L'espace des lignes contient les vecteurs $\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ et $\begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ et le noyau contient le

vecteur $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

c) L'équation $Ax = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ admet une solution et $A^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

d) A est non nulle et chaque ligne est orthogonale à chaque colonne.

e) La somme des vecteurs lignes est le vecteur formé que de 1 et la somme des colonnes est le vecteur nul.

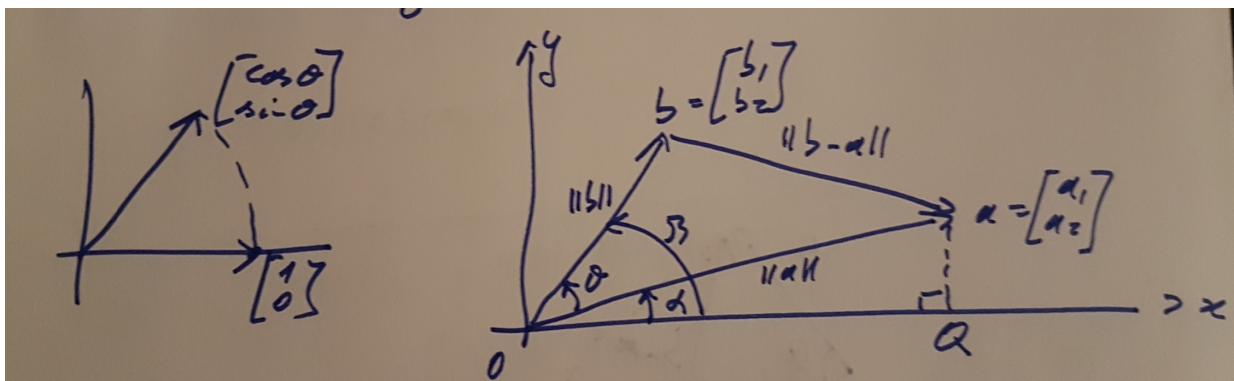
2.3.14. Exercice. — Construire une matrice A de $\mathcal{M}_n(\mathbb{R})$ sans terme nul et dont les vecteurs colonnes sont deux à deux orthogonaux. Calculer $A^T A$.

2.3.15. Exercice. — Soit A une matrice inversible de $\mathcal{M}_n(\mathbb{R})$. La première colonne de A^{-1} est orthogonale à l'espace engendré par quelles lignes de A ?

Décomposition matricielle QR

§ 1 Cosinus et projection sur une droite

3.1.1. Cosinus d'un angle. —



Par définition, on a

$$\begin{aligned}\sin \alpha &= \frac{a_2}{\|\mathbf{a}\|} & \cos \alpha &= \frac{a_1}{\|\mathbf{a}\|} \\ \sin \beta &= \frac{b_2}{\|\mathbf{b}\|} & \cos \beta &= \frac{b_1}{\|\mathbf{b}\|}\end{aligned}$$

On a

$$\theta = \beta - \alpha$$

et donc

$$\cos \theta = \cos \beta \cos \alpha + \sin \beta \sin \alpha$$

soit

$$\cos \theta = \frac{a_1 b_1 + a_2 b_2}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|} = \frac{\mathbf{a}^\top \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}.$$

Le *cosinus* de l'angle θ formé par les deux vecteurs $\mathbf{a}$ et $\mathbf{b}$ est donc donné par l'expression vectorielle suivante :

$$\cos \theta = \frac{\mathbf{a}^\top \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}.$$

Remarques. Notons que l'on a toujours $\mathbf{a}^\top \mathbf{b} = \mathbf{b}^\top \mathbf{a}$.

Par ailleurs, la valeur du cosinus ne dépend que de la direction des vecteurs $\mathbf{a}$ et $\mathbf{b}$. En effet, supposons que $\mathbf{b}' = 2\mathbf{b}$, alors

$$\frac{\mathbf{a}^\top \mathbf{b}'}{\|\mathbf{a}\| \cdot \|\mathbf{b}'\|} = \frac{\mathbf{a}^\top 2\mathbf{b}}{\|\mathbf{a}\| \cdot \|2\mathbf{b}\|} = \frac{\mathbf{a}^\top \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}.$$

3.1.2 Proposition. — Pour tous vecteurs $\mathbf{a}$ et $\mathbf{b}$, on a :

$$\|\mathbf{b} - \mathbf{a}\|^2 = \|\mathbf{b}\|^2 + \|\mathbf{a}\|^2 - 2\|\mathbf{b}\| \cdot \|\mathbf{a}\| \cos \theta.$$

Preuve. Si les vecteurs $\mathbf{a}$ et $\mathbf{b}$ sont orthogonaux, alors $\mathbf{a}^\top \mathbf{b} = 0$ et $\cos \theta = 0$. L'égalité est donc le théorème de Pythagore.

Si les vecteurs $\mathbf{a}$ et $\mathbf{b}$ ne sont pas orthogonaux, on a

$$\|\mathbf{b} - \mathbf{a}\|^2 = (\mathbf{b} - \mathbf{a})^\top (\mathbf{b} - \mathbf{a}) = \mathbf{b}^\top \mathbf{b} - \mathbf{b}^\top \mathbf{a} - \mathbf{a}^\top \mathbf{b} + \mathbf{a}^\top \mathbf{a}.$$

Par ailleurs, on a

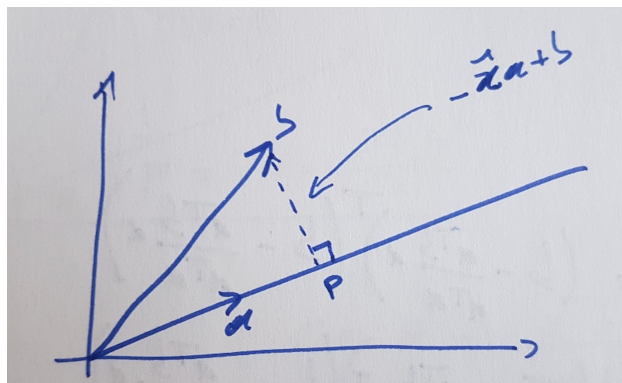
$$\|\mathbf{b}\|^2 + \|\mathbf{a}\|^2 - 2\|\mathbf{b}\| \cdot \|\mathbf{a}\| \cos \theta = \mathbf{b}^\top \mathbf{b} + \mathbf{a}^\top \mathbf{a} - 2\|\mathbf{b}\| \cdot \|\mathbf{a}\| \cos \theta.$$

On a ainsi

$$\|\mathbf{b} - \mathbf{a}\|^2 - \|\mathbf{b}\|^2 - \|\mathbf{a}\|^2 + 2\|\mathbf{b}\| \cdot \|\mathbf{a}\| \cos \theta = -2\mathbf{a}^\top \mathbf{b} + 2\|\mathbf{b}\| \cdot \|\mathbf{a}\| \cos \theta.$$

L'égalité à démontrer découle ainsi de la formule du cosinus $\cos \theta = \frac{\mathbf{a}^\top \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}$. $\square$

3.1.3. Projection sur une droite. — On veut déterminer la *projection orthogonale* $\mathbf{p}$ du vecteur $\mathbf{b}$ sur la droite portée par le vecteur $\mathbf{a}$:



On a

$$\mathbf{p} = \hat{x}\mathbf{a},$$

où $\hat{x}$ est un réel. Déterminons ce réel.

Par construction, les vecteurs $(\mathbf{b} - \hat{x}\mathbf{a})$ et $\mathbf{a}$ sont orthogonaux. On a donc

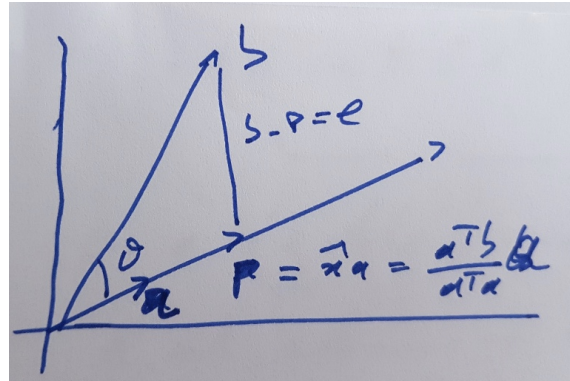
$$\mathbf{a}^\top (\mathbf{b} - \hat{x}\mathbf{a}) = 0$$

Ainsi

$$\hat{x} = \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}}.$$

Cette formule nous permet de définir la *projection orthogonale* du vecteur $\mathbf{b}$ sur la droite portée par le vecteur $\mathbf{a}$ comme

$$\mathbf{p} = \hat{x}\mathbf{a} = \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} \mathbf{a}.$$



On a $\cos \theta = \frac{\mathbf{a}^\top \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}$ et

$$\|\mathbf{e}\|^2 = \|\mathbf{b} - \mathbf{p}\|^2 \geq 0.$$

Ainsi,

$$\left\| \mathbf{b} - \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} \mathbf{a} \right\|^2 = \mathbf{b}^\top \mathbf{b} - 2 \frac{(\mathbf{a}^\top \mathbf{b})^2}{\mathbf{a}^\top \mathbf{a}} + \left(\frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} \right)^2 \mathbf{a}^\top \mathbf{a} \geq 0$$

D'où

$$\frac{(\mathbf{b}^\top \mathbf{b})(\mathbf{a}^\top \mathbf{a}) - (\mathbf{a}^\top \mathbf{b})^2}{\mathbf{a}^\top \mathbf{a}} \geq 0$$

Ainsi,

$$(\mathbf{b}^\top \mathbf{b})(\mathbf{a}^\top \mathbf{a}) \geq (\mathbf{a}^\top \mathbf{b})^2$$

qui démontre le résultat suivant :

3.1.4 Proposition (Inégalité de Schwarz). — Pour tous vecteurs $\mathbf{a}$ et $\mathbf{b}$, on a

$$|\mathbf{a}^\top \mathbf{b}| \leq \|\mathbf{a}\| \cdot \|\mathbf{b}\|.$$

3.1.5. Remarque. — Comme $\frac{\mathbf{a}^\top \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|} = \cos \theta$, l'inégalité de Schwarz est équivalente à

$$|\cos \theta| \leq 1.$$

C'est-à-dire

$$-1 \leq \cos \theta \leq 1.$$

On a l'égalité, si et seulement si le vecteur $\mathbf{a}$ est colinéaire au vecteur $\mathbf{b}$, c'est-à-dire si $\cos \theta = 1$ soit $\theta = 0$.

3.1.6. Exemple. — Considérons la projection $\mathbf{p}$ du vecteur $\mathbf{b}$ sur la droite portée par le vecteur $\mathbf{a}$, avec

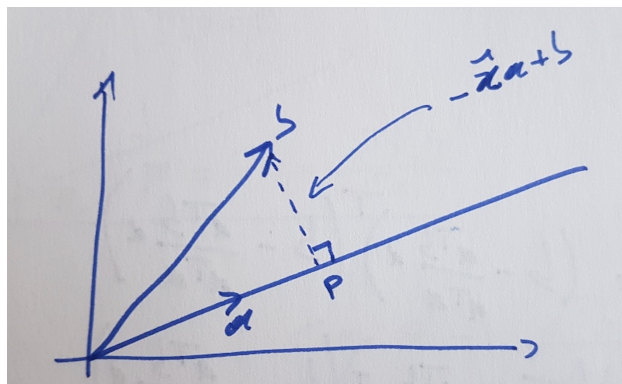
$$\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \text{et} \quad \mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

On a

$$\hat{x} = \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} = \frac{6}{3} = 2.$$

Ainsi $\mathbf{p} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$.

3.1.7. Matrice de la projection. — On considère la situation



On a

$$\mathbf{p} = \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} \mathbf{a}.$$

avec $\frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} \in \mathbb{R}$. Donc

$$\mathbf{p} = \mathbf{a} \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} = \frac{\mathbf{a} \mathbf{a}^\top}{\mathbf{a}^\top \mathbf{a}} \mathbf{b}.$$

Ainsi, la matrice de la projection du vecteur $\mathbf{b}$ sur la droite portée par $\mathbf{a}$ est

$$\mathbf{P} = \frac{\mathbf{a} \mathbf{a}^\top}{\mathbf{a}^\top \mathbf{a}}.$$

3.1.8. Exemple. — La projection sur le vecteur $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ est ainsi

$$\mathbf{P} = \frac{\mathbf{a} \mathbf{a}^\top}{\mathbf{a}^\top \mathbf{a}} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

On vérifie que

$$\frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.$$

3.1.9. Remarques. — On notera que la matrice $\mathbf{P}$ de la projection vérifie

i) $\mathbf{P}^2 = \mathbf{P}$,

ii) $\mathbf{P}$ est symétrique, c'est-à-dire que $\mathbf{P} = \mathbf{P}^\top$,

iii) $\text{Im}(\mathbf{P}) = \text{Vect}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$

iv) $\text{Ker}(\mathbf{P})$ est le plan perpendiculaire au vecteur $\mathbf{a}$ engendré par les vecteurs $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ et

$$\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

v) $\text{rg}(\mathbf{P}) = 1$

§ 2 Projections orthogonale et méthode des moindres carrés

Nous avons vu que l'équation $\mathbf{Ax} = \mathbf{b}$ possède des solutions si $\mathbf{b}$ est dans l'espace colonnes($\mathbf{A}$). Sinon, le système d'équation $\mathbf{Ax} = \mathbf{b}$ est inconsistent et il n'y a pas de solution.

Par exemple, le système

$$\begin{cases} 2x = b_1 \\ 3x = b_2 \\ 4x = b_3 \end{cases}$$

possède des solutions uniquement si $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ est dans le sous espace $\text{Vect}\left(\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}\right)$.

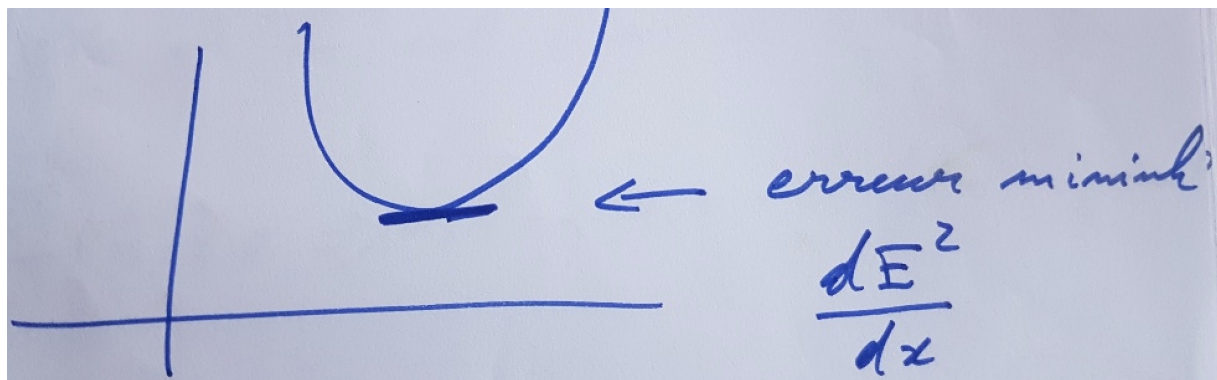
En pratique, on doit pouvoir trouver une solution (approchée) dans tous les cas. Pour cela, on peut essayer de résoudre le système en ignorant une partie des équations. L'idée est de minimiser l'erreur que l'on fait en faisant cela.

Considérons l'erreur

$$E^2 = (2x - b_1)^2 + (3x - b_2)^2 + (4x - b_3)^2.$$

S'il existe une solution exacte, on a $E = 0$.

Lorsque $\mathbf{b}$ n'est pas dans $\text{Vect}\left(\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}\right)$, le graphe de $E^2(x)$ est une parabole :



On a

$$\frac{dE^2}{dx} = 2[2(2x - b_1) + 3(3x - b_2) + 4(4x - b_3)].$$

On cherche à résoudre

$$\frac{dE^2}{dx} = 0.$$

3.2.1. Solution des moindres carrés. — La *solution des moindres carrés* de l'équation

$$\mathbf{ax} = \mathbf{b}$$

est

$$\hat{x} = \frac{2b_1 + 3b_2 + 4b_3}{2^2 + 3^2 + 4^2} = \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}}.$$

Dans le cas général dans $\mathbb{R}^m$, résoudre $\mathbf{ax} = \mathbf{b}$ en minimisant

$$E(x)^2 = \|\mathbf{ax} - \mathbf{b}\|^2 = (a_1x - b_1)^2 + \dots + (a_mx - b_m)^2.$$

La dérivée de $E(x)^2$ est 0 en $\hat{x}$. On a

$$\frac{dE^2}{dx}(\hat{x}) = 0$$

implique

$$(a_1\hat{x} - b_1)a_1 + \dots + (a_m\hat{x} - b_m)a_m = 0.$$

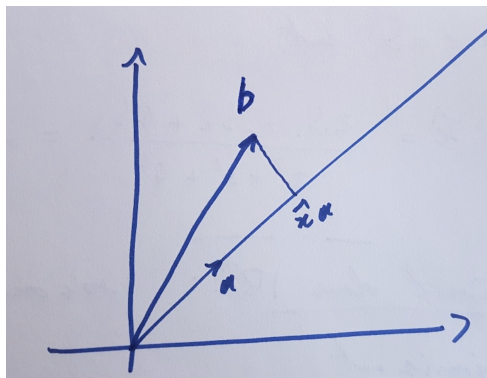
Donc

$$\hat{x} = \frac{a_1b_1 + \dots + a_mb_m}{a_1^2 + \dots + a_m^2}$$

La solution des moindres carrés de l'équation $\mathbf{ax} = \mathbf{b}$ dans $\mathbb{R}^m$ est

$$\hat{x} = \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}}.$$

L'erreur commise est $\mathbf{b} - \hat{x}\mathbf{a}$:



$\mathbf{b} - \hat{x}\mathbf{a}$ doit être orthogonal à $\mathbf{a}$. En effet,

$$\mathbf{a}^\top (\mathbf{b} - \hat{x}\mathbf{a}) = \mathbf{a}^\top \mathbf{b} - \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}} \mathbf{a}^\top \mathbf{a} = 0.$$

Cela confirme ainsi notre définition de $\hat{x}$.

3.2.2. Méthode des moindres carrés dans le cas de plusieurs variables. — On considère $\mathbf{A}$ dans $\mathcal{M}_{m,n}(\mathbb{R})$:

$$\begin{bmatrix} a_1^1 & a_1^2 & \cdots & a_1^n \\ a_2^1 & a_2^2 & \cdots & a_2^n \\ \vdots & \vdots & & \vdots \\ a_m^1 & a_m^2 & \cdots & a_m^n \end{bmatrix}$$

et $\mathbf{b}$ dans $\mathbb{R}^m$.

On cherche à résoudre le système

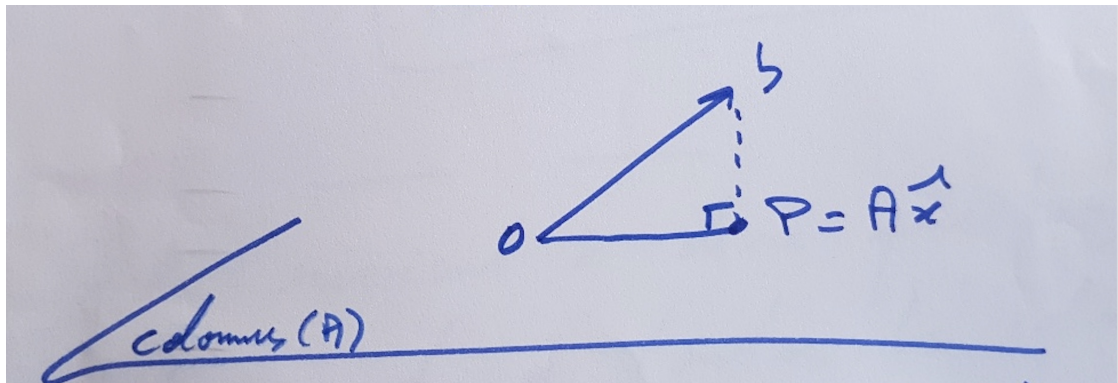
$$\mathbf{A}\mathbf{x} = \mathbf{b}.$$

On peut avoir $m > n$. Le système est inconsistant si $\mathbf{b}$ n'est pas dans l'espace colonnes($\mathbf{A}$).

Le problème consiste à trouver une solution $\hat{\mathbf{x}}$ qui minimise l'erreur :

$$E = \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}\|.$$

Cela correspond à la distance de $\mathbf{b}$ à l'espace des colonnes colonnes($\mathbf{A}$) :



Trouver $\hat{\mathbf{x}}$ qui minimise l'erreur, c'est trouver $\mathbf{A}\hat{\mathbf{x}}$ dans colonnes($\mathbf{A}$) qui est le plus proche de $\mathbf{b}$.

En fait $\mathbf{p} = \mathbf{A}\hat{\mathbf{x}}$ doit être la projection orthogonale de $\mathbf{b}$ sur colonnes($\mathbf{A}$), c'est-à-dire que

$$\mathbf{e} = (\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) \perp \text{colonnes}(\mathbf{A}).$$

Or on a

$$(\text{colonnes}(\mathbf{A}))^\perp = \text{Ker}(\mathbf{A}^\top).$$

Donc le vecteur d'erreur $\mathbf{e} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$ doit vérifier

$$\mathbf{A}^\top(\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) = \mathbf{0}.$$

Soit

$$\mathbf{A}^\top \mathbf{A} \hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}.$$

Ce qui est équivalent à dire que le vecteur $\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$ est orthogonal à toutes les colonnes de la matrice $\mathbf{A}$. Dans les faits, cela revient à multiplier chaque membre de l'équation $\mathbf{A}\mathbf{x} = \mathbf{b}$ par la matrice $\mathbf{A}^\top$:

$$\mathbf{A}^\top \mathbf{A} \mathbf{x} = \mathbf{A}^\top \mathbf{b}.$$

3.2.3. En résumé. —

- i) Lorsque le système $\mathbf{Ax} = \mathbf{b}$ est inconsistant (pas de solution), la *solution des moindres carrés* $\hat{\mathbf{x}}$ qui minimise $\|\mathbf{Ax} - \mathbf{b}\|^2$ vérifie

$$\mathbf{A}^\top \mathbf{A} \hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}.$$

- ii) Par ailleurs, la matrice $\mathbf{A}^\top \mathbf{A}$ est inversible lorsque les colonnes de la matrice $\mathbf{A}$ sont linéairement indépendantes. La meilleure approximation de $\hat{\mathbf{x}}$ est

$$\hat{\mathbf{x}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}.$$

- iii) La projection du vecteur $\mathbf{b}$ sur l'espace colonnes($\mathbf{A}$) est le vecteur

$$\mathbf{p} = \mathbf{A} \hat{\mathbf{x}} = \mathbf{A} (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}.$$

3.2.4. Exemple. — Considérons

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}, \quad \text{et} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

Le système $\mathbf{Ax} = \mathbf{b}$ ne possède pas de solution :

$$\left| \begin{array}{l} x + 2y = 4 \\ x + 3y = 5 \\ 0 + 0 = 6 \end{array} \right.$$

On a

$$\mathbf{A}^\top \mathbf{A} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}$$

et

$$\hat{\mathbf{x}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b} = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

La « meilleure façon » de résoudre le système $\mathbf{Ax} = \mathbf{b}$ est donc

$$\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

L'erreur commise est donc

$$\mathbf{e} = \mathbf{b} - \mathbf{A} \hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}.$$

3.2.5. Exercice. — Déterminer la projection orthogonale du vecteur $\mathbf{b} = \begin{bmatrix} 2 \\ 4 \\ 4 \end{bmatrix}$ sur la droite

portée par le vecteur $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

3.2.6. Exercice. — Déterminer dans $\mathbb{R}^n$ l'angle du vecteur de composantes $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ avec les axes. Déterminer la matrice $\mathbf{P}$ de la projection orthogonale sur ce vecteur.

3.2.7. Exercice. — Utiliser un vecteur approprié dans l'inégalité de Schwarz pour montrer que

$$(a_1 + \dots + a_n)^2 \leq n(a_1^2 + \dots + a_n^2)$$

Quand a-t-on l'égalité ?

3.2.8. Exercice. — Soit $\mathbf{a}$ un vecteur de $\mathbb{R}^n$. On pose

$$\mathbf{P} = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top \mathbf{a}}.$$

Montrer que $\mathbf{P}^2 = \mathbf{P}$. Montrer que la trace de la matrice $\mathbf{P}$ est égale à 1.

3.2.9. Exercice. — On considère le vecteur $\mathbf{a} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

1. Déterminer la matrice $\mathbf{P}_1$ de la projection orthogonale sur $\mathbf{a}$ et la matrice $\mathbf{P}_2$ de la projection orthogonale sur un vecteur perpendiculaire à $\mathbf{a}$.
2. Calculer $\mathbf{P}_1 + \mathbf{P}_2$ et $\mathbf{P}_1\mathbf{P}_2$. Expliquer les résultats obtenus.

3.2.10. Exercice. — Déterminer la matrice de la projection orthogonale de $\mathbb{R}^3$ sur l'intersection des plans d'équation $x + y + t = 0$ et $x - t = 0$.

3.2.11. Exercice. — Soit $\mathbf{A}$ une matrice. Démontrer que si $\mathbf{A}\mathbf{A}^\top = \mathbf{A}^\top \mathbf{A}$ alors pour tout $\mathbf{x}$

$$\|\mathbf{A}^\top \mathbf{x}\| = \|\mathbf{A}\mathbf{x}\|$$

3.2.12. Exercice. — Déterminer la solution $\hat{x}$ des moindres carrés des équations $3x = 10$ et $4x = 5$. Vérifier que l'erreur commise est un vecteur orthogonal au vecteur $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

3.2.13. Exercice. — Déterminer la solution $\hat{x}$ des moindres carrés l'équation $\mathbf{A}\mathbf{x} = \mathbf{b}$ et déterminer $\mathbf{p} = \mathbf{A}\hat{\mathbf{x}}$ pour

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Vérifier que l'erreur $\mathbf{b} - \mathbf{p}$ est perpendiculaire aux vecteurs colonnes de $\mathbf{A}$.

3.2.14. Exercice. — Déterminer la matrice $\mathbf{P}$ de la projection sur le sous-espace de $\mathbb{R}^3$ engendré par les vecteurs $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ et $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$.

3.2.15. Exercice. — Soit V le sous-espace de $\mathbb{R}^4$ engendré par les vecteurs

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Déterminer :

1. une base du supplémentaire orthogonal $V^\perp$
2. la matrice $\mathbf{P}$ de la projection orthogonale sur V
3. le vecteur de V le plus proche du vecteur $\mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$ de $V^\perp$.

3.2.16. Exercice. — On pose

$$\mathbf{a} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad \text{et} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

des vecteurs de $\mathbb{R}^m$. Soit $\hat{\mathbf{x}}$ la solution des moindres carrés des m équations $\mathbf{a}\mathbf{x} = \mathbf{b}$.

1. Résoudre l'équation $\mathbf{a}^\top \mathbf{a} \hat{\mathbf{x}} = \mathbf{a}^\top \mathbf{b}$ pour montrer que $\hat{\mathbf{x}}$ est la moyenne des réels $b_1, \dots, b_m$.
2. Calculer le vecteur $\mathbf{e} = \mathbf{b} - \mathbf{a}\hat{\mathbf{x}}$, puis $\|\mathbf{e}\|^2$ (appelée *variance* des b_i) et $\|\mathbf{e}\|$ (appelé *écart-type* des b_i).
3. Pour $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}$ vérifier que le vecteur $\mathbf{p} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$ est perpendiculaire au vecteur $\mathbf{e}$ et déterminer la matrice $\mathbf{P}$ de la projection sur la droite portée par le vecteur $\mathbf{a}$.

§ 3 Bases orthonormées

3.3.1. Définitions. — Des vecteurs $\mathbf{q}_1, \dots, \mathbf{q}_n$ de $\mathbb{R}^m$ sont dit *orthonormaux*, ou la famille $(\mathbf{q}_1, \dots, \mathbf{q}_n)$ est dite *orthonormale*, s'ils sont de norme égale à 1 et orthogonaux deux-à-deux. C'est-à-dire, si

$$\mathbf{q}_i^\top \mathbf{q}_j = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j. \end{cases}$$

Une base de $\mathbb{R}^n$ est dite *orthonormée*, ou *base orthogonale*, si elle est formée d'une famille orthonormale de vecteurs.

3.3.2. Les matrices orthogonales. — Une matrice $\mathbf{Q}$ de $\mathcal{M}_n(\mathbb{K})$ est dite *orthogonale* si ses vecteurs colonnes sont orthonormaux.

3.3.3. Exemples. — La base canonique $\mathcal{Can}$ de $\mathbb{R}^n$ est orthogonale. De la même façon, la matrice identité $\mathbf{1}_n$ est orthogonale.

3.3.4 Proposition. — Une matrice $\mathbf{Q}$ est orthogonale si et seulement si $\mathbf{Q}^\top \mathbf{Q} = \mathbf{1}_n$.

Preuve. On a

$$\begin{bmatrix} - & \mathbf{q}_1^\top & - \\ - & \mathbf{q}_2^\top & - \\ & \vdots & \\ - & \mathbf{q}_n^\top & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^\top \mathbf{q}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{q}_2^\top \mathbf{q}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{q}_n^\top \mathbf{q}_n \end{bmatrix} = \mathbf{1}_n$$

□

L'inverse d'une matrice orthogonale est sa transposée. En effet, si $\mathbf{Q}^\top \mathbf{Q} = \mathbf{1}_n$ alors

$$\mathbf{Q}^\top = \mathbf{Q}^{-1}.$$

3.3.5. Exemple. — La matrice de la rotation d'angle θ

$$\mathbf{Q} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

son inverse est la rotation d'angle $-\theta$:

$$\mathbf{Q}^{-1} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

3.3.6. Exemple. — Une matrice de permutation est une matrice qui a les mêmes colonnes que la matrice identité, peu importe l'ordre :

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \mathbf{P}^\top = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

3.3.7 Proposition. — La multiplication par une matrice orthonogonale $\mathbf{Q}$

i) préserve la norme, c'est-à-dire

$$\|\mathbf{Q}\mathbf{x}\| = \|\mathbf{x}\|,$$

pour tout $\mathbf{x} \in \mathbb{R}^n$,

ii) préserve le produit scalaire, c'est-à-dire

$$(\mathbf{Q}\mathbf{x})^\top (\mathbf{Q}\mathbf{y}) = \mathbf{x}^\top \mathbf{y}$$

pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Preuve. Montrons i), pour tout $\mathbf{x} \in \mathbb{R}^n$, on a

$$(\mathbf{Q}\mathbf{x})^\top (\mathbf{Q}\mathbf{x}) = \mathbf{x}^\top \mathbf{Q}^\top \mathbf{Q} \mathbf{x} = \mathbf{x}^\top \mathbf{x}.$$

D'où $\|\mathbf{Q}\mathbf{x}\| = \|\mathbf{x}\|$. □

3.3.8. Décomposition d'un vecteur dans une base orthonormée. — Soit $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base de $\mathbb{R}^n$. Tout vecteur $\mathbf{b}$ de $\mathbb{R}^n$ se décompose de façon unique en

$$\mathbf{b} = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n,$$

où les $x_1, \dots, x_n$ dans $\mathbb{R}$ sont les coordonnées du vecteur $\mathbf{b}$.

Nous allons montrer que trouver les coordonnées d'un tel vecteur dans une base orthonormée est simple.

3.3.9 Proposition. — Soit $(\mathbf{q}_1, \dots, \mathbf{q}_n)$ une base orthonormée de $\mathbb{R}^n$, alors pour tout vecteur $\mathbf{b}$ dans $\mathbb{R}^n$, on a

$$\mathbf{b} = (\mathbf{q}_1^\top \mathbf{b}) \mathbf{q}_1 + (\mathbf{q}_2^\top \mathbf{b}) \mathbf{q}_2 + \dots + (\mathbf{q}_n^\top \mathbf{b}) \mathbf{q}_n.$$

Preuve. Calculons les x_i dans une décomposition unique

$$\mathbf{b} = x_1 \mathbf{q}_1 + \dots + x_n \mathbf{q}_n.$$

Soit $1 \leq i \leq n$, on a

$$\mathbf{q}_i^\top \mathbf{b} = \mathbf{q}_i^\top (x_1 \mathbf{q}_1) + \dots + \mathbf{q}_i^\top (x_i \mathbf{q}_i) + \dots + \mathbf{q}_i^\top (x_n \mathbf{q}_n).$$

Et donc

$$\mathbf{q}_i^\top \mathbf{b} = x_1 \mathbf{q}_i^\top \mathbf{q}_1 + \dots + x_i \mathbf{q}_i^\top \mathbf{q}_i + \dots + x_n \mathbf{q}_i^\top \mathbf{q}_n = x_i.$$

□

3.3.10. Remarque. — Déterminer les scalaires $x_1, \dots, x_n$ tels que

$$\mathbf{b} = x_1 \mathbf{q}_1 + \dots + x_n \mathbf{q}_n,$$

revient à résoudre le système linéaire $\mathbf{Q}\mathbf{x} = \mathbf{b}$, avec

$$\mathbf{Q} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & & \mathbf{q}_n \\ | & | & & | \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

De plus, résoudre ce système revient à calculer la transposée de la matrice $\mathbf{Q}$. En effet, on a $\mathbf{Q}^\top \mathbf{Q} = \mathbf{1}_n$, donc

$$\mathbf{x} = \mathbf{Q}^\top \mathbf{b} = \begin{bmatrix} \mathbf{q}_1^\top \mathbf{b} \\ \vdots \\ \mathbf{q}_n^\top \mathbf{b} \end{bmatrix}.$$

3.3.11 Proposition. — Les lignes d'une matrice carrée forment une famille orthonormale si, et seulement si, ses colonnes forment une famille orthonormale.

Preuve. On a $\mathbf{Q}^\top \mathbf{Q} = \mathbf{1}_n$ si et seulement si $\mathbf{Q}\mathbf{Q}^\top = \mathbf{1}_n$. □

3.3.12. Exemple. —

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

3.3.13. Le cas des matrices rectangulaires. — On dit qu'une matrice $\mathbf{Q}$ de $\mathcal{M}_{m,n}(\mathbb{K})$ est *orthogonale* si ses vecteurs colonnes forment une famille orthonormale.

Comme dans le cas des matrices carrées de $\mathcal{M}_n(\mathbb{K})$, c'est équivalent à dire que $\mathbf{Q}^\top \mathbf{Q} = \mathbf{1}_n$.

3.3.14. . — On considère le système linéaire suivant

$$\mathbf{Q}\mathbf{x} = \mathbf{b}$$

où

$$\mathbf{Q} = \begin{bmatrix} | & & | \\ \mathbf{q}_1 & \cdots & \mathbf{q}_n \\ | & & | \end{bmatrix}$$

est une matrice orthogonale de $\mathcal{M}_{m,n}(\mathbb{K})$, avec $m > n$. On a ainsi un système linéaire avec plus d'équations que d'inconnues :

$$\begin{cases} q_1^1 x_1 + \dots + q_1^n x_n = b_1 \\ q_2^1 x_1 + \dots + q_2^n x_n = b_2 \\ \vdots \\ q_m^1 x_1 + \dots + q_m^n x_n = b_m \end{cases}$$

Ce système n'a pas de solution pour tout vecteur $\mathbf{b}$ de $\mathbb{R}^n$ en général.

3.3.15. Résolution par la méthode des moindres carrés. — On peut résoudre l'équation $\mathbf{Q}\mathbf{x} = \mathbf{b}$ par la méthode des moindres carrés de la façon suivante.

On a

$$\mathbf{Q}^\top \mathbf{Q} \hat{\mathbf{x}} = \mathbf{Q}^\top \mathbf{b}$$

Comme $\mathbf{Q}^\top \mathbf{Q} = \mathbf{1}_n$, on a

$$\hat{\mathbf{x}} = \mathbf{Q}^\top \mathbf{b}$$

Autrement dit, $\hat{x}_i = \mathbf{Q}_i^\top \mathbf{b}$ pour tout $1 \leq i \leq n$.

Ainsi, la solution est exacte si la matrice $\mathbf{Q}$ est carrée et c'est la solution des moindres carrés lorsque la matrice $\mathbf{Q}$ est rectangulaire.

3.3.16. Remarques. — La projection du vecteur $\mathbf{b}$ sur l'espace des colonnes $\text{colonnes}(\mathbf{Q})$ est

$$\mathbf{p} = \mathbf{Q}\hat{\mathbf{x}} = (\mathbf{q}_1^\top \mathbf{b})\mathbf{q}_1 + \dots + (\mathbf{q}_n^\top \mathbf{b})\mathbf{q}_n.$$

On a $\mathbf{p} = \mathbf{Q}\mathbf{Q}^\top \mathbf{b}$. La matrice de la projection est donc

$$\mathbf{P} = \mathbf{Q}\mathbf{Q}^\top.$$

§ 4 Orthonormalisation de Gram-Schmidt

Soient $\mathbf{a}$, $\mathbf{b}$ et $\mathbf{c}$ des vecteurs de $\mathbb{R}^n$. Si $\mathbf{a}$, $\mathbf{b}$ et $\mathbf{c}$ sont orthonormaux, la projection d'un vecteur $\mathbf{v}$ sur le sous-espace vectoriel $\text{Vect}(\mathbf{a}, \mathbf{b}, \mathbf{c})$ engendré par les trois vecteurs $\mathbf{a}$, $\mathbf{b}$ et $\mathbf{c}$ est

$$\mathbf{p} = (\mathbf{a}^\top \mathbf{v})\mathbf{a} + (\mathbf{b}^\top \mathbf{v})\mathbf{b} + (\mathbf{c}^\top \mathbf{v})\mathbf{c}.$$

Que faire si les trois vecteurs $\mathbf{a}$, $\mathbf{b}$ et $\mathbf{c}$ ne forment pas une famille orthonormale ?

On va les rendre orthonormaux avec le procédé de Gram-Schmidt.

On pose

$$\mathbf{q}_1 = \frac{\mathbf{a}}{\|\mathbf{a}\|}$$

On a $\|\mathbf{q}_1\| = 1$. On pose

$$\mathbf{b}' = \mathbf{b} - (\mathbf{q}_1^\top \mathbf{b})\mathbf{q}_1, \quad \text{et} \quad \mathbf{q}_2 = \frac{\mathbf{b}'}{\|\mathbf{b}'\|}$$

On fait de même pour le troisième vecteur. On pose :

$$\mathbf{c}' = \mathbf{c} - (\mathbf{q}_1^\top \mathbf{c})\mathbf{q}_1 - (\mathbf{q}_2^\top \mathbf{c})\mathbf{q}_2, \quad \text{et} \quad \mathbf{q}_3 = \frac{\mathbf{c}'}{\|\mathbf{c}'\|}$$

On obtient ainsi une famille $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$ qui est une base orthonormée de l'espace $\text{Vect}(\mathbf{a}, \mathbf{b}, \mathbf{c})$. Ce processus est appelé le processus d'orthonormalisation de Gram-Schmidt.

3.4.1. Factorisation QR. — On considère une matrice

$$\mathbf{A} = \begin{bmatrix} | & | & | \\ \mathbf{a} & \mathbf{b} & \mathbf{c} \\ | & | & | \end{bmatrix}$$

On construit la matrice

$$\mathbf{Q} = \begin{bmatrix} | & | & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ | & | & | \end{bmatrix}$$

où

$$\begin{aligned} \mathbf{a} &= (\mathbf{q}_1^\top \mathbf{a})\mathbf{q}_1 \\ \mathbf{b} &= (\mathbf{q}_1^\top \mathbf{b})\mathbf{q}_1 + (\mathbf{q}_2^\top \mathbf{b})\mathbf{q}_2 \\ \mathbf{c} &= (\mathbf{q}_1^\top \mathbf{c})\mathbf{q}_1 + (\mathbf{q}_2^\top \mathbf{c})\mathbf{q}_2 + (\mathbf{q}_3^\top \mathbf{c})\mathbf{q}_3. \end{aligned}$$

On montre que $\mathbf{A} = \mathbf{QR}$. En effet

$$\mathbf{A} = \begin{bmatrix} | & | & | \\ \mathbf{a} & \mathbf{b} & \mathbf{c} \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^\top \mathbf{a} & \mathbf{q}_1^\top \mathbf{b} & \mathbf{q}_1^\top \mathbf{c} \\ 0 & \mathbf{q}_2^\top \mathbf{b} & \mathbf{q}_2^\top \mathbf{c} \\ 0 & 0 & \mathbf{q}_3^\top \mathbf{c} \end{bmatrix}$$

La matrice dernière matrice est une matrice triangulaire supérieure notée

$$\mathbf{R} = \begin{bmatrix} \mathbf{q}_1^\top \mathbf{a} & \mathbf{q}_1^\top \mathbf{b} & \mathbf{q}_1^\top \mathbf{c} \\ 0 & \mathbf{q}_2^\top \mathbf{b} & \mathbf{q}_2^\top \mathbf{c} \\ 0 & 0 & \mathbf{q}_3^\top \mathbf{c} \end{bmatrix}.$$

La décomposition précédente est appelée la décomposition QR de la matrice $\mathbf{A}$.

3.4.2. Cas général. — Toute matrice $\mathbf{A}$ de $\mathcal{M}_{m,n}(\mathbb{K})$ peut être factorisée en

$$\mathbf{A} = \mathbf{QR},$$

où la matrice $\mathbf{Q}$ est orthogonale et la matrice $\mathbf{R}$ est triangulaire supérieure et inversible.

3.4.3. Retour sur la méthode des moindres carrés. — On considère le système linéaire

$$\mathbf{A}\mathbf{x} = \mathbf{b}.$$

La solution des moindres carrés de ce système est le vecteur $\hat{\mathbf{x}}$ solution du système

$$\mathbf{A}^\top \mathbf{A} \hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}.$$

Si on a la décomposition QR

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

alors

$$\mathbf{A}^\top \mathbf{A} = \mathbf{R}^\top \mathbf{Q}^\top \mathbf{Q} \mathbf{R} = \mathbf{R}^\top \mathbf{R}.$$

Ainsi, le système $\mathbf{A}^\top \mathbf{A} \hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}$ est équivalent au système

$$\mathbf{R}^\top \mathbf{R} \hat{\mathbf{x}} = \mathbf{R}^\top \mathbf{Q}^\top \mathbf{b}$$

Autrement dit, cela revient à résoudre le système triangulaire suivant

$$\mathbf{R} \hat{\mathbf{x}} = \mathbf{Q}^\top \mathbf{b}.$$

§ 5 Exercices

3.5.1. Exercice. — On considère les quatre vecteurs suivants de $\mathbb{R}^3$:

$$\mathbf{a}_1 = \begin{bmatrix} 2/3 \\ 2/3 \\ -1/3 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} -1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}.$$

1. Calculer les projections du vecteur $\mathbf{b}$ sur les vecteurs normés $\mathbf{a}_1$ et $\mathbf{a}_2$ puis sa projection sur le plan engendré par les vecteurs $\mathbf{a}_1$ et $\mathbf{a}_2$.
2. Calculer de même la projection de $\mathbf{b}$ sur $\mathbf{a}_3$. Expliquer pourquoi la matrice

$$\mathbf{a}_1 \mathbf{a}_1^\top + \mathbf{a}_2 \mathbf{a}_2^\top + \mathbf{a}_3 \mathbf{a}_3^\top$$

est la matrice identité.

3.5.2. Exercice. — Démontrer que si $\mathbf{u}$ est un vecteur unitaire alors la matrice

$$\mathbf{Q} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^\top$$

est une matrice symétrique orthogonale. Calculer cette matrice pour le vecteur $\mathbf{u} = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix}$.

3.5.3. Exercice. — Démontrer que toute matrice orthogonale et triangulaire supérieure est nécessairement diagonale.

3.5.4. Exercice. — Appliquer l'algorithme de Gram-Schmidt aux trois vecteurs suivants

$$\mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

et écrire le résultat sous la forme de la décomposition $\mathbf{A} = \mathbf{QR}$ où $\mathbf{A} = [\mathbf{a} \ \mathbf{b} \ \mathbf{c}]$.

3.5.5. Exercice. — Construire une base orthonormale à partir des trois vecteurs suivants

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

3.5.6. Exercice. — Déterminer une base orthonormale $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ telle que $\mathbf{q}_1$ et $\mathbf{q}_2$ engendrent l'espace des vecteurs colonnes de la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}$$

Déterminer la solution des moindres carrés de l'équation $\mathbf{Ax} = \mathbf{b}$ avec $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}$.

3.5.7. Exercice. — Utiliser une décomposition $\mathbf{A} = \mathbf{QR}$ de la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 1 \end{bmatrix}$$

pour trouver la solution des moindres carrés de l'équation $\mathbf{Ax} = \mathbf{b}$ avec $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

3.5.8. Exercice. — Si $\mathbf{A} = \mathbf{QR}$ déterminer une expression simple de la matrice $\mathbf{P}$ de la projection sur l'espace des vecteurs colonnes de $\mathbf{A}$.

3.5.9. Exercice. — Déterminer une base orthonormale du sous-espace engendré par les trois vecteurs

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

3.5.10. Exercice. — Partant des vecteurs suivants

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix},$$

utiliser la méthode de Gram-Schmidt pour déterminer une base orthonormale du plan d'équation $x_1 + x_2 + x_3 = 0$.

3.5.11. Exercice. — Les deux assertions suivantes sont-elles vraies ?

- 1 Si Q est une matrice orthogonale alors Q^{-1} est orthogonale.
- 2 Si Q est une matrice 3 est une matrice dont les deux vecteurs colonnes forment une base orthonormale alors pour tout vecteur x on a $\| Qx \| = \| x \|$.

Méthodes itératives pour la résolution des systèmes linéaires

§ 1 Les méthodes itératives

Les méthodes de résolutions directes de systèmes linéaires de la forme $Ax = b$ vues dans les chapitres précédents, comme la méthode LU ou la méthode QR, permettent d'obtenir des solutions exactes, mais leur calculs peuvent être fastidieux pour les systèmes de grande taille. Ces méthodes présentent également des problèmes dans leur implémentation sur machine.

L'idée des méthodes itératives est de ne pas chercher une solution exacte, mais une solution approchée.

4.1.1. Problème. — Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$, avec $\mathbb{K} = \mathbb{R}$, ou $\mathbb{K} = \mathbb{C}$. On suppose que A est inversible. Pour un vecteur $b \in \mathbb{R}^n$, on cherche une approximation d'un vecteur x de $\mathbb{R}^n$ solution de l'équation

$$Ax = b.$$

4.1.2. Méthode. — On pose

$$A = M - N,$$

où M est une matrice « facile » à inverser. C'est-à-dire, une matrice diagonale, triangulaire, orthogonale. Le choix de la matrice M est guidée par la nature de la matrice A .

Résoudre le système $Ax = b$ est alors équivalent à résoudre le système

$$Mx = Nx + b.$$

Si la matrice M est facile à inverser, cela revient à résoudre

$$x = M^{-1}Nx + M^{-1}b. \quad (4.1)$$

On définit la fonction

$$F(x) = M^{-1}Nx + M^{-1}b$$

et on cherche x point fixe de cette fonction, c'est-à-dire vérifiant $x = F(x)$.

4.1.3. Matrice de Jacobi. — La matrice $M^{-1}N$ dans la décomposition (4.1) est appelée la *matrice de Jacobi* de la décomposition $A = M - N$.

4.1.4. Méthode itérative. — On construit une suite $((x^{(k)})_{k \geq 0})$ de façon récursive par l'algorithme suivant

$$\begin{cases} x^{(0)} \text{ fixé} \\ x^{(k+1)} = F(x^{(k)}) = M^{-1}Nx^{(k)} + M^{-1}b, \quad \forall k \geq 0. \end{cases} \quad (4.2)$$

Si la suite $(x^{(k)})_{k \geq 0}$ converge, alors sa limite, notée x^∞ , doit satisfaire

$$x^\infty = F(x^\infty),$$

car F est une fonction affine et continue.

Donc

$$x^\infty = M^{-1}Nx^\infty + M^{-1}b.$$

Soit

$$Ax^\infty = b.$$

La solution $x^\infty = A^{-1}b$ est appelé *point fixe*, ou *point stationnaire* de l'algorithme (4.2).

4.1.5. Critère de convergence. — Soit $A = M - N$ une matrice de $\mathcal{M}_n(\mathbb{R})$, où M est inversible. L'algorithme itératif (4.2) converge si pour tout vecteur b de $\mathbb{R}^n$, tout vecteur $x^{(0)}$ de $\mathbb{R}^n$, la suite $(x^{(k)})_{k \geq 0}$ converge vers la solution $x = A^{-1}b$ dans $\mathbb{R}^n$. C'est-à-dire

$$\lim_{k \rightarrow +\infty} \|x^{(k)} - x\| = 0.$$

Notons que le critère de convergence porte sur la décomposition $A = M - N$ et non pas sur $X^{(0)}$ et b .

4.1.6. Question de la définition du critère. — Comment établir un critère sur les matrices M et N pour assurer la convergence ?

Notons $e^{(k)}$ l'erreur entre la solution approchée $x^{(k)}$ et la solution x :

$$e^{(k)} = x^{(k)} - x.$$

On ne connaît pas x . L'idée est de trouver un critère pour que

$$\lim_{k \rightarrow +\infty} e^{(k)} = 0.$$

$$\begin{aligned} e^{(k+1)} &= x^{(k+1)} - x \\ &= M^{-1}Nx^{(k)} + M^{-1}b - x \\ &= M^{-1}Nx^{(k)} + M^{-1}b - M^{-1}(Nx + b) \\ &= M^{-1}N(x^{(k)} - x). \end{aligned}$$

Ainsi,

$$e^{(k+1)} = M^{-1}Ne^{(k)}.$$

Posons $B = M^{-1}N$. On a

$$e^{(k+1)} = Be^{(k)}$$

pour tout $k \geq 0$. Soit

$$e^{(k+1)} = B^{k+1}e^{(0)}.$$

L'algorithme itératif (4.2) converge si pour tout vecteur $e^{(0)}$ de $\mathbb{R}^n$, on a

$$\lim_{k \rightarrow +\infty} B^k e^{(0)} = 0.$$

4.1.7. Rayon spectral et critère de convergence. — Soit $\mathbf{B}$ une matrice de $\mathcal{M}_n(\mathbb{R})$.

Rappelons qu'un réel λ est *valeur propre* de la matrice $\mathbf{B}$ si

$$\det(\mathbf{B} - \lambda \mathbf{1}_n) = 0.$$

On note $\text{Sp}(\mathbf{B})$ l'ensemble des valeurs propres de la matrice $\mathbf{B}$. On notera

$$\text{Sp}_{\mathbb{C}}(\mathbf{B}) = \{\lambda \in \mathbb{C} \mid \det(\mathbf{B} - \lambda \mathbf{1}_n) = 0\}.$$

Le *rayon spectral* de la matrice $\mathbf{B}$ est défini par

$$\rho(\mathbf{B}) = \max\{|\lambda| \mid \lambda \in \text{Sp}_{\mathbb{C}}(\mathbf{B})\}.$$

4.1.8 Théorème. — Soit $\mathbf{B}$ une matrice de $\mathcal{M}_n(\mathbb{R})$. Les trois assertions suivantes sont équivalentes

i) Pour tout vecteur $\mathbf{v}$ de $\mathbb{R}^n$, $\lim_{k \rightarrow +\infty} \mathbf{B}^k \mathbf{v} = \mathbf{0}$,

ii) $\rho(\mathbf{B}) < 1$,

iii) $\|\mathbf{B}\| < 1$.

§ 2 Méthode de Jacobi

On considère la décomposition

$$\mathbf{A} = \mathbf{D} - \mathbf{E} - \mathbf{F},$$

où

i) $\mathbf{D}$ est la diagonale de la matrice $\mathbf{A}$,

ii) $-\mathbf{E}$ est la partie de la matrice $\mathbf{A}$ sous la diagonale,

iii) $-\mathbf{F}$ est la partie de la matrice $\mathbf{A}$ sur la diagonale.

C'est-à-dire

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} a_1^1 & a_1^2 & \cdots & a_1^n \\ a_2^1 & a_2^2 & \cdots & a_2^n \\ \vdots & \vdots & \ddots & \vdots \\ a_m^1 & a_m^2 & \cdots & a_m^n \end{bmatrix} & \mathbf{D} &= \begin{bmatrix} a_1^1 & 0 & \cdots & 0 \\ 0 & a_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_m^n \end{bmatrix} \\ -\mathbf{E} &= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ a_2^1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_m^1 & \cdots & a_m^{n-1} & 0 \end{bmatrix} & -\mathbf{F} &= \begin{bmatrix} 0 & a_1^2 & \cdots & a_1^n \\ 0 & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & a_m^{n-1} \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \end{aligned}$$

On suppose que la diagonale ne possède pas d'éléments nuls.

On pose $\mathbf{M} = \mathbf{D}$ et $\mathbf{N} = \mathbf{E} + \mathbf{F}$.

L'algorithme itératif (4.2) s'écrit

$$\begin{cases} \mathbf{x}^{(0)} \in \mathbb{R}^n \text{ fixé} \\ \mathbf{D}\mathbf{x}^{(k+1)} = (\mathbf{E} + \mathbf{F})\mathbf{x}^{(k)} + \mathbf{b}, \quad \forall k \geq 0. \end{cases} \quad (4.3)$$

Autrement dit avec $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

$$\left| \begin{array}{l} \mathbf{x}^{(0)} \in \mathbb{R}^n \text{ fixé} \\ a_i^j x_i^{k+1} = -\sum_{i < j} a_i^j x_j^{(k)} - \sum_{j < i} a_i^j x_j^{(k)} + b_i \quad \forall k \geq 0, \forall i \in \{1, \dots, n\}. \end{array} \right. \quad (4.4)$$

On a

$$x_i^{(k+1)} = \frac{1}{a_i} \left(b_i - \sum_{j=1, j \neq i}^n a_i^j x_j^{(k)} \right).$$

Ainsi, à partir de $\mathbf{x}^{(0)}$, on calcul les $\mathbf{x}^{(k)}$ avec la solution précédente.

4.2.1. Méthode de Jacobi. — On fixe un paramètre $\epsilon > 0$ et $\mathbf{x}_0 \in \mathbb{R}^n$.

Poser $\mathbf{x}^{(0)} = \mathbf{x}_0$, $\epsilon^{(0)} = 2\epsilon$, $k = 0$.

tant que $\epsilon^{(k)} \geq \epsilon$ **faire**

$$\left| \begin{array}{l} \mathbf{x}^{(k+1)} = \mathbf{D}^{-1}((\mathbf{E} + \mathbf{F})\mathbf{x}^{(k)} + \mathbf{b}) \\ \epsilon^{(k+1)} := \|\mathbf{A}\mathbf{x}^{(k+1)} - \mathbf{b}\| \\ k := k + 1 \end{array} \right.$$

fin

Sorties : le vecteur $\mathbf{x}^{(k)}$

4.2.2 Théorème. — Soit $\mathbf{A}$ une matrice de $\mathcal{M}_n(\mathbb{R})$ de diagonale dominante, i.e.,

$$|a_i^i| > \sum_{j \neq i} a_i^j, \quad \forall i \in \{1, \dots, n\}.$$

Alors pour tout $\mathbf{x}^{(0)}$ la méthode de Jacobi converge vers la solution de $\mathbf{A}\mathbf{x} = \mathbf{b}$.

4.2.3. Exemple. — On considère la matrice

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

On pose

$$\mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

La matrice $\mathbf{A}$ est à diagonale dominante, donc la méthode de Jacobi converge.

On a

$$\mathbf{B} = \mathbf{D}^{-1}(\mathbf{E} + \mathbf{F}) = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix}.$$

$$\text{Posons } \mathbf{x}^{(k)} = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} \text{ et } \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

On a

$$\begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix} \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

On a $\lim_{k \rightarrow +\infty} \mathbf{B}^k = \mathbf{0}$.

4.2.4. Exemple. — On considère la matrice

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

On pose

$$\mathbf{D} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & -1 & 0 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 0 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Posons $\mathbf{x}^{(k)} = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{bmatrix}$ et $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

On a

$$\mathbf{x}^{(k+1)} = \mathbf{D}^{-1}(\mathbf{E} + \mathbf{F})\mathbf{x}^{(k)} + \mathbf{D}^{-1}\mathbf{b}.$$

On a

$$\mathbf{B} = \mathbf{D}^{-1}(\mathbf{E} + \mathbf{F}) = -\frac{1}{3} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

et

$$\mathbf{D}^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}.$$

$$\mathbf{x}^{(k+1)} = -\frac{1}{3} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \mathbf{x}^{(k)} + \frac{1}{3} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Le polynôme caractéristique de la matrice $\mathbf{B}$ est

$$P_{\mathbf{B}} = (x + \frac{2}{3})(x - \frac{1}{3})^2.$$

On a $\text{Sp}(\mathbf{B}) = \{-\frac{2}{3}, \frac{1}{3}\}$ et $\rho(\mathbf{B}) = \frac{2}{3}$.

§ 3 Méthode de Gauss-Seidel

On décompose

$$\mathbf{A} = \mathbf{D} - \mathbf{E} - \mathbf{F}.$$

On pose $\mathbf{M} = \mathbf{D} - \mathbf{E}$ et $\mathbf{N} = \mathbf{F}$.

L'algorithme itératif (4.2) s'écrit

$$\left| \begin{array}{l} \mathbf{x}^{(0)} \in \mathbb{R}^n \text{ fixé} \\ (\mathbf{D} - \mathbf{E})\mathbf{x}^{(k+1)} = \mathbf{F}\mathbf{x}^{(k)} + \mathbf{b}, \quad \forall k \geq 0. \end{array} \right. \quad (4.5)$$

La matrice $\mathbf{D} - \mathbf{E}$ est triangulaire inférieure facile à inverser.

Autrement dit avec $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

$$\left| \begin{array}{l} \mathbf{x}^{(0)} \in \mathbb{R}^n \text{ fixé} \\ a_i^i x_i^{(k+1)} = -\sum_{i < j} a_i^j x_j^{(k)} - \sum_{j < i, 1 \leq i \leq n, k \geq 0} a_i^j x_j^{(k)} + b_i \quad \forall k \geq 0, \forall i \in \{1, \dots, n\}. \end{array} \right. \quad (4.6)$$

4.3.1. Méthode de Gauss-Seidel. — On fixe un paramètre réel $\epsilon > 0$ et un vecteur $\mathbf{x}_0$ de $\mathbb{R}^n$.

Poser $\mathbf{x}^{(0)} = \mathbf{x}_0$, $\epsilon^{(0)} = 2\epsilon$, $k = 0$.

tant que $\epsilon^{(k)} \geq \epsilon$ **faire**

$\mathbf{x}^{(k+1)} := (\mathbf{D} - \mathbf{E})^{-1}(\mathbf{F}\mathbf{x}^{(k)} + \mathbf{b})$
 $\epsilon^{(k+1)} := \|\mathbf{A}\mathbf{x}^{(k+1)} - \mathbf{b}\|$
 $k := k + 1$

fin

Sorties : le vecteur $\mathbf{x}^{(k)}$

4.3.2 Théorème. — Soit $\mathbf{A}$ une matrice de $\mathcal{M}_n(\mathbb{R})$. Si la matrice $\mathbf{A}$ est symétrique et définie positive, alors pour tout vecteur $\mathbf{x}^{(0)}$ de $\mathbb{R}^n$, la méthode de Gauss-Seidel converge vers la solution du système $\mathbf{A}\mathbf{x} = \mathbf{b}$.

§ 4 Exercices

4.4.1. Exercice. — On considère la matrice

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

1. Déterminer sa matrice de Jacobi et les valeurs propres de celle-ci.
2. Déterminer sa matrice de Gauss-Seidel et les valeurs propres de celle-ci.

4.4.2. Exercice. — Soit $\mathbf{A}$ la matrice

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

1. Déterminer sa matrice de Jacobi et ses valeurs propres λ_1 et λ_2 .
2. Déterminer sa matrice de Gauss-Seidel et ses valeurs propres μ_1 et μ_2 .
3. Comparer $\lambda_{\max}$ et $\mu_{\max}$.

4.4.3. Exercice. — On pose

$$\mathbf{A} = \begin{bmatrix} 10 & 1 \\ 2 & 10 \end{bmatrix}$$

et $\mathbf{b} = \begin{bmatrix} 11 \\ 12 \end{bmatrix}$.

1. Donner la solution de l'équation $\mathbf{A}\mathbf{x} = \mathbf{b}$.
2. Calculer le rayon spectral de la matrice de Jacobi de $\mathbf{A}$.
3. Utiliser la méthode de Jacobi pour trouver une valeur approchée de la solution.

4.4.4. Exercice. — On pose

$$\mathbf{A} = \begin{bmatrix} 1 & 10 \\ 10 & 2 \end{bmatrix} \quad \text{et} \quad \mathbf{b} = \begin{bmatrix} 11 \\ 12 \end{bmatrix}.$$

1. Donner la solution de l'équation $\mathbf{A}\mathbf{x} = \mathbf{b}$.
2. Calculer le rayon spectral de la matrice de Jacobi de $\mathbf{A}$.
3. La méthode de Jacobi donne-t-elle une solution approchée ?

4.4.5. Exercice. — On pose

$$\mathbf{A} = \begin{bmatrix} 10 & 1 \\ 2 & 10 \end{bmatrix} \quad \text{et} \quad \mathbf{b} = \begin{bmatrix} 11 \\ 12 \end{bmatrix}.$$

1. Donner la solution de l'équation $\mathbf{Ax} = \mathbf{b}$.
2. Calculer le rayon spectral de la matrice de Gauss-Seidel de $\mathbf{A}$.
3. Utiliser la méthode de Gauss-Seidel pour trouver une valeur approchée de la solution.

4.4.6. Exercice. — On pose

$$\mathbf{A} = \begin{bmatrix} 1 & 10 \\ 10 & 2 \end{bmatrix} \quad \text{et} \quad \mathbf{b} = \begin{bmatrix} 11 \\ 12 \end{bmatrix}.$$

1. Donner la solution de l'équation $\mathbf{Ax} = \mathbf{b}$.
2. Calculer le rayon spectral de la matrice de Gauss-Seidel de $\mathbf{A}$.
3. La méthode de Gauss-Seidel donne-t-elle une solution approchée ?

Valeurs propres et vecteurs propres

Dans une première partie de ce chapitre, nous rappelons les notions de valeurs propres et de vecteurs propres abordées dans le cadre du cours d'AMALA A.

§ 1 Résoudre $A\mathbf{x} = \lambda\mathbf{x}$

5.1.1. Migration de population. — On étudie la répartition de la population entre deux zones géographiques d'un même pays, disons entre une région nord et une région sud. Chaque année, un mouvement de population entre ces deux zones s'opère de la façon suivante :

- 80% de la population des régions du nord part vivre dans les régions du sud,
- 70% de la population des régions du sud rejoint les régions nords.

On suppose que la population totale du pays reste constante d'une année à l'autre, ainsi 20% des habitants des régions du nord restent dans le nord et 30% des habitants des régions du sud restent dans le sud. Le mouvement de population est schématisé par la figure 5.1

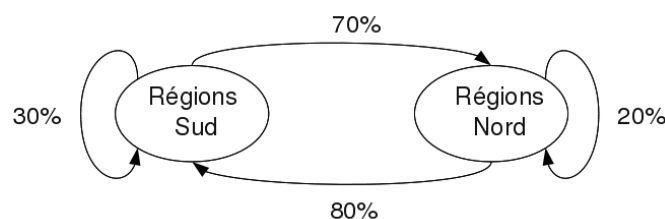


FIGURE 5.1 – Mouvement de population entre deux zones géographiques

Étant donné une répartition initiale, l'année $k = 0$, de la population entre les deux zones, quelle sera la répartition de population la k -ième année ? Est-ce qu'au bout d'un certain temps, toute la population va se retrouver en zone sud, ou bien le nombre d'habitants va-t-il se stabiliser dans chaque zone ?

5.1.2. Modélisation du problème. — Notons n_k le nombre d'habitants des régions du nord la k -ième année et s_k le nombre d'habitants des régions du sud la k -ième année. L'évolution des

deux populations est régie par le système d'équations couplées suivant

$$\begin{cases} n_{k+1} = 0.2n_k + 0.7s_k \\ s_{k+1} = 0.8n_k + 0.3s_k \end{cases}$$

qui s'écrit matriciellement sous la forme :

$$\begin{bmatrix} n_{k+1} \\ s_{k+1} \end{bmatrix} = \mathbf{A} \begin{bmatrix} n_k \\ s_k \end{bmatrix},$$

avec

$$\mathbf{A} = \begin{bmatrix} 0.2 & 0.7 \\ 0.8 & 0.3 \end{bmatrix}$$

et où le vecteur $\mathbf{p}(k) = \begin{bmatrix} n_k \\ s_k \end{bmatrix}$ exprime la répartition de la population au nord et au sud l'année k . Par exemple, si l'année $k = 0$, on a un million d'habitants au nord et dix mille habitants au sud, *i.e.*, $\mathbf{p}(0) = \begin{bmatrix} 10^6 \\ 10^4 \end{bmatrix}$, alors l'année suivante, à $k = 1$, il y aura

$$\begin{aligned} \mathbf{p}(1) &= \begin{bmatrix} n_1 \\ s_1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} 10^6 \\ 10^4 \end{bmatrix} \\ &= \begin{bmatrix} 207000 \\ 803000 \end{bmatrix}. \end{aligned}$$

Pour connaître la population dans chacune des régions la k -ième année, on peut appliquer la formule suivante, obtenue par récurrence sur le nombre d'années k :

$$\begin{bmatrix} n_k \\ s_k \end{bmatrix} = \mathbf{A}^k \begin{bmatrix} n_0 \\ s_0 \end{bmatrix}.$$

Voici les valeurs du vecteur $\mathbf{p}(k)$ pour quelques valeurs de k :

$$\begin{aligned} \mathbf{p}(1) &= \begin{bmatrix} 207000 \\ 803000 \end{bmatrix}, & \mathbf{p}(2) &= \begin{bmatrix} 603500 \\ 406500 \end{bmatrix}, & \mathbf{p}(3) &= \begin{bmatrix} 405250 \\ 604750 \end{bmatrix}, \\ \mathbf{p}(4) &= \begin{bmatrix} 504375 \\ 505625 \end{bmatrix}, & \mathbf{p}(5) &= \begin{bmatrix} 454812.5 \\ 555187.5 \end{bmatrix}, & \mathbf{p}(6) &= \begin{bmatrix} 479593.75 \\ 530406.25 \end{bmatrix}, \\ \mathbf{p}(7) &= \begin{bmatrix} 467203.12 \\ 542796.87 \end{bmatrix}, & \mathbf{p}(8) &= \begin{bmatrix} 473398.43 \\ 536601.56 \end{bmatrix}, & \mathbf{p}(30) &= \begin{bmatrix} 471333.33 \\ 538666.66 \end{bmatrix}. \end{aligned} \quad (5.1)$$

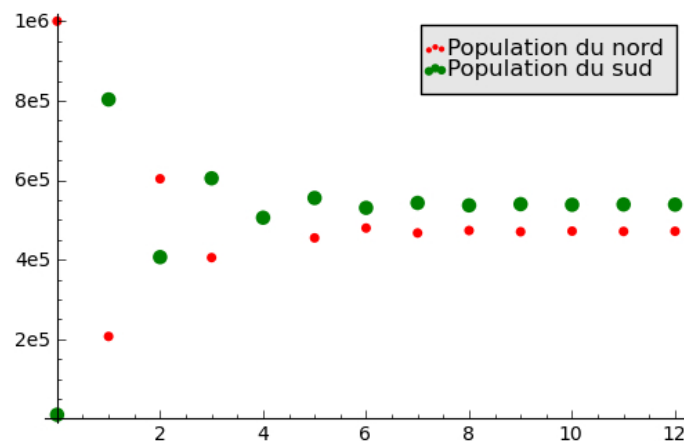


FIGURE 5.2 – Évolution de la population entre le nord et le sud les 12 premières années

Nous allons montrer que le calcul des puissances k -ième successives de la matrice $\mathbf{A}$ peut se faire à l'aide d'un simple calcul de valeurs propres de la matrice $\mathbf{A}$. Cette approche permet d'éviter les multiplications de la matrice $\mathbf{A}$ par elle-même, cette opération étant coûteuse en nombre d'opérations. Cette approche va nous permettre aussi de mieux comprendre l'évolution du système, ici de l'évolution de la population dans chaque zone géographique, en fonction du vecteur de répartition initial $\mathbf{p}(0)$.

5.1.3. Calcul des valeurs propres. — Nous allons donner dans ce chapitre une méthode de calcul des valeurs propres pour une matrice : les valeurs propres de la matrice $\mathbf{A}$ sont les solutions λ de l'équation suivante :

$$\det(\mathbf{A} - \lambda \mathbf{1}_2) = 0.$$

Soit

$$\begin{vmatrix} 0.2 - \lambda & 0.7 \\ 0.8 & 0.3 - \lambda \end{vmatrix} = \lambda^2 - 0.5\lambda - 0.5 = (\lambda - 1)(\lambda + 0.5) = 0.$$

La matrice $\mathbf{A}$ possède ainsi deux valeurs propres, $\lambda = 1$ et $\lambda = -0.5$, qui sont les racines du polynôme $\lambda^2 - 0.5\lambda - 0.5$.

Si λ est valeur propre, comme $\det(\mathbf{A} - \lambda \mathbf{1}_2) = 0$, la matrice $\mathbf{A} - \lambda \mathbf{1}_2$ admet un noyau non trivial. Autrement dit, il existe des vecteurs $\mathbf{x}$ solutions de l'équation suivante

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

Ces vecteurs sont dans la même direction que $\mathbf{A}\mathbf{x}$, on les appelle *vecteurs propres* associés à la valeur propre λ . La valeur propre indique si les vecteurs propres sont laissés inchangés, s'ils sont étirés, réduits ou encore inversés. Dans notre exemple, les vecteurs propres associés à la valeur propre $\lambda = 1$ restent inchangés, ils vérifient $\mathbf{A}\mathbf{x} = \mathbf{x}$. Ceux associés à la valeur propre $\lambda = -0.5$ sont réduits et inversés : $\mathbf{A}\mathbf{x} = -0.5\mathbf{x}$.

5.1.4. Les sous-espaces propres. — Lorsque l'on calcule les puissances de la matrice $\mathbf{A}$, les valeurs propres sont élevées à la puissance, alors que les vecteurs propres restent inchangés. En effet, pour tout vecteur propre $\mathbf{x}$ associé à la valeur propre -0.5 , on a

$$\mathbf{A}^2\mathbf{x} = \mathbf{A}(\mathbf{A}\mathbf{x}) = -0.5\mathbf{A}\mathbf{x} = (-0.5)^2\mathbf{x},$$

ainsi, pour tout entier k ,

$$\mathbf{A}^k\mathbf{x} = (-0.5)^k\mathbf{x}.$$

Pour un vecteur propre $\mathbf{x}$ associé à la valeur propre 1, on a $\mathbf{A}^k\mathbf{x} = \mathbf{x}$, pour tout entier k .

Par suite, en décomposant les vecteurs dans une base de vecteurs propres, on pourra alors calculer les puissances de $\mathbf{A}$ simplement en calculant les puissances des valeurs propres.

L'ensemble des vecteurs propres associés à la valeur propre -0.5 est défini par

$$E_{-0.5} = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{A}\mathbf{x} = -0.5\mathbf{x}\}.$$

Donc $E_{-0.5}$ est formé des vecteurs $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ dont les composantes satisfont le système d'équations suivant :

$$\begin{cases} -0.5x_1 = 0.2x_1 + 0.7x_2 \\ -0.5x_2 = 0.8x_1 + 0.3x_2 \end{cases}$$

Ainsi, $E_{-0.5}$ est l'espace vectoriel de dimension 1 engendré par le vecteur $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. De la même façon, on montre que l'espace vectoriel formé des vecteurs propres associés à la valeur propre 1, défini par

$$E_1 = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{A}\mathbf{x} = \mathbf{x}\}.$$

est engendré par le vecteur $\begin{bmatrix} 0.875 \\ 1 \end{bmatrix}$. Ces deux vecteurs forment une base de $\mathbb{R}^2$. En particulier, étant donné le vecteur initial $\begin{bmatrix} n_0 \\ s_0 \end{bmatrix} = \begin{bmatrix} 10^6 \\ 10^4 \end{bmatrix}$, il existe deux réels uniques β_1 et β_2 tels que

$$\begin{bmatrix} n_0 \\ s_0 \end{bmatrix} = \beta_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta_2 \begin{bmatrix} 0.875 \\ 1 \end{bmatrix}.$$

De cette équation, on déduit les deux réels β_1 et β_2 , on a

$$\beta_1 = \frac{1646.10^3}{3} \quad \text{et} \quad \beta_2 = \frac{1616.10^3}{3}.$$

On peut alors exprimer le vecteur $\begin{bmatrix} n_k \\ s_k \end{bmatrix}$, on a

$$\begin{aligned} \begin{bmatrix} n_k \\ s_k \end{bmatrix} &= \mathbf{A}^k \begin{bmatrix} n_0 \\ s_0 \end{bmatrix} \\ &= \beta_1 \mathbf{A}^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta_2 \mathbf{A}^k \begin{bmatrix} 0.875 \\ 1 \end{bmatrix} \\ &= \beta_1 (-0.5)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta_2 \begin{bmatrix} 0.875 \\ 1 \end{bmatrix} \end{aligned}$$

D'où

$$\begin{bmatrix} n_k \\ s_k \end{bmatrix} = \frac{1646.10^3}{3} (-0.5)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{1616.10^3}{3} \begin{bmatrix} 0.875 \\ 1 \end{bmatrix}$$

On a ainsi résolu le problème du calcul de la répartition de la population la k -ième année sans avoir à faire le produit de la matrice $\mathbf{A}$ par elle-même k -fois. Si k tend vers l'infini, on a de façon immédiate que

$$\lim_{k \rightarrow +\infty} \begin{bmatrix} n_k \\ s_k \end{bmatrix} = \frac{1616.10^3}{3} \begin{bmatrix} 0.875 \\ 1 \end{bmatrix} = \begin{bmatrix} 471333, 33 \\ 538666, 66 \end{bmatrix}$$

Nous retrouvons les valeurs numériques obtenues en (5.1).

§ 2 Valeurs propres et espaces propres

5.2.1. Définitions. — Soit $\mathbf{A}$ une matrice de $\mathcal{M}_n(\mathbb{K})$. Un scalaire λ dans $\mathbb{K}$ est appelé *valeur propre* de $\mathbf{A}$, si l'une des conditions équivalentes suivantes est vérifiée :

- i) $\text{Ker}(\mathbf{A} - \lambda \mathbf{1}_n) \neq \{0\}$,
- ii) $\det(\mathbf{A} - \lambda \mathbf{1}_n) = 0$,
- iii) il existe un vecteur non nul $\mathbf{x}$ de $\mathbb{K}^n$, solution de l'équation

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

L'équivalence des assertions **i)**, **ii)** est immédiate. L'assertion **iii)** est une simple reformulation de **i)**

Un vecteur $\mathbf{x}$ vérifiant l'assertion **iii)** est appelé *vecteur propre* de $\mathbf{A}$ associé à la valeur propre λ . Le sous ensemble de $\mathbb{K}$ formé des valeurs propres de $\mathbf{A}$ est appelé le *spectre* sur $\mathbb{K}$ de la matrice $\mathbf{A}$ et sera noté dans la suite $\text{Sp}_{\mathbb{K}}(\mathbf{A})$, ou $\text{Sp}(\mathbf{A})$ s'il n'y a pas de confusion possible avec le corps de base.

5.2.2. Exemple. — Considérons la matrice suivante de $\mathcal{M}_2(\mathbb{R})$:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Alors, on a

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

ainsi le vecteur $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ est un vecteur propre de $\mathbf{A}$ associé à la valeur propre 1. Par ailleurs, on a

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -1 \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

le vecteur $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ est un vecteur propre de $\mathbf{A}$ associé à la valeur propre -1 .

5.2.3. Sous-espaces propres. —

5.2.4 Proposition. — Soit $\mathbf{A}$ une matrice de $\mathcal{M}_n(\mathbb{K})$ et soit λ une valeur propre de $\mathbf{A}$. L'ensemble des vecteurs propres associés à λ

$$E_{\lambda} = \{\mathbf{x} \in \mathbb{K}^n \mid \mathbf{A}\mathbf{x} = \lambda\mathbf{x}\}$$

est un sous-espace vectoriel de $\mathbb{K}^n$. De plus, le sous-espace E_{λ} est stable par $\mathbf{A}$, *i.e.*, le vecteur $\mathbf{A}\mathbf{x}$ appartient à E_{λ} , pour tout vecteur $\mathbf{x}$ de E_{λ} .

Preuve. On vérifie que l'ensemble E_{λ} contient le vecteur nul et que si $\mathbf{x}$ et $\mathbf{y}$ sont deux vecteurs de E_{λ} et α et β deux scalaires, alors

$$\begin{aligned} \mathbf{A}(\alpha\mathbf{x} + \beta\mathbf{y}) &= \alpha\mathbf{A}\mathbf{x} + \beta\mathbf{A}\mathbf{y}, \\ &= \alpha\lambda\mathbf{x} + \beta\lambda\mathbf{y}, \\ &= \lambda(\alpha\mathbf{x} + \beta\mathbf{y}). \end{aligned}$$

Montrons la stabilité du sous-espace E_{λ} par $\mathbf{A}$. Si $\mathbf{x} \in E_{\lambda}$, on a

$$\mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}(\lambda\mathbf{x}) = \lambda\mathbf{A}\mathbf{x},$$

donc $\mathbf{A}\mathbf{x} \in E_{\lambda}$. $\square$

5.2.5. Définition. — L'espace E_λ est appelé le *sous-espace propre* de la matrice $\mathbf{A}$ associé à la valeur propre λ .

5.2.6. Exemple. — Nous avons montré que la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

admet pour valeur propres -1 et 1 . Les sous-espaces propres associés sont

$$E_{-1} = \text{Vect}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right), \quad E_1 = \text{Vect}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right).$$

5.2.7. Somme de sous-espaces propres. — Soit $\mathbf{A}$ une matrice de $\mathcal{M}_n(\mathbb{K})$ admettant deux valeurs propres distinctes λ_1 et λ_2 . Les sous-espaces propres E_{λ_1} et E_{λ_2} sont alors en somme directe. En effet, si $\mathbf{x}$ est un vecteur propre de $\mathbf{A}$ associé aux deux valeurs propres λ_1 et λ_2 . On a

$$\mathbf{A}\mathbf{x} = \lambda_1\mathbf{x} \quad \text{et} \quad \mathbf{A}\mathbf{x} = \lambda_2\mathbf{x}.$$

D'où

$$(\lambda_1 - \lambda_2)\mathbf{x} = \mathbf{0}.$$

Comme les valeurs propres λ_1 et λ_2 sont supposées distinctes, on en déduit que $\mathbf{x} = \mathbf{0}$. On a donc

$$E_{\lambda_1} \cap E_{\lambda_2} = \{\mathbf{0}\}.$$

C'est ainsi que E_{λ_1} et E_{λ_2} sont en somme directe. Plus généralement on a :

5.2.8 Proposition. — Soient $\lambda_1, \dots, \lambda_p$ des valeurs propres d'une matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{K})$, distinctes deux à deux. Les sous-espaces propres $E_{\lambda_1}, \dots, E_{\lambda_p}$ sont en somme directe.

Preuve. On procède par récurrence sur le nombre p de valeurs propres distinctes. Si $p = 1$, il n'y a rien à démontrer. Supposons que les sous-espaces $E_{\lambda_1}, \dots, E_{\lambda_p}$ soient en somme directe. C'est équivalent à

$$E_{\lambda_1} \cap E_{\lambda_2} = \{\mathbf{0}\}, \quad (E_{\lambda_1} + E_{\lambda_2}) \cap E_{\lambda_3} = \{\mathbf{0}\}, \quad \dots, \quad (E_{\lambda_1} + E_{\lambda_2} + \dots + E_{\lambda_{p-1}}) \cap E_{\lambda_p} = \{\mathbf{0}\}.$$

Pour montrer que les sous-espaces $E_{\lambda_1}, \dots, E_{\lambda_p}, E_{\lambda_{p+1}}$ sont en somme directe, il suffit de montrer que

$$(E_{\lambda_1} + E_{\lambda_2} + \dots + E_{\lambda_p}) \cap E_{\lambda_{p+1}} = \{\mathbf{0}\}.$$

Soit $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_p$ un vecteur de $(E_{\lambda_1} + E_{\lambda_2} + \dots + E_{\lambda_p}) \cap E_{\lambda_{p+1}}$, où $\mathbf{x}_k \in E_{\lambda_k}$, pour tout $k \in \llbracket 1, p \rrbracket$. Comme $\mathbf{A}\mathbf{x}_k = \lambda_k\mathbf{x}_k$, on a

$$\mathbf{A}\mathbf{x} = \lambda_1\mathbf{x}_1 + \dots + \lambda_p\mathbf{x}_p.$$

Par ailleurs, $\mathbf{x}$ est un vecteur propre associé à λ_{p+1} , d'où

$$\mathbf{A}\mathbf{x} = \lambda_{p+1}\mathbf{x} = \lambda_{p+1}\mathbf{x}_1 + \dots + \lambda_{p+1}\mathbf{x}_p.$$

Par suite,

$$\mathbf{0} = (\lambda_1 - \lambda_{p+1})\mathbf{x}_1 + \cdots + (\lambda_p - \lambda_{p+1})\mathbf{x}_p.$$

Les sous-espaces $E_{\lambda_1}, \dots, E_{\lambda_p}$ étant en somme directe, $\mathbf{0} = \mathbf{0}_{E_{\lambda_1}} + \cdots + \mathbf{0}_{E_{\lambda_p}}$ est la seule décomposition de $\mathbf{0}$. D'où, pour tout $k \in \llbracket 1, p \rrbracket$,

$$(\lambda_k - \lambda_{p+1})\mathbf{x}_k = \mathbf{0}_{E_{\lambda_k}},$$

Comme les valeurs propres λ_i sont deux à deux distinctes, tous les vecteurs $\mathbf{x}_k$ sont nuls, par suite le vecteur $\mathbf{x}$ est nul. Ainsi, le sous-espace $(E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_p}) \cap E_{\lambda_{p+1}}$ est trivial et les sous-espaces $E_{\lambda_1}, \dots, E_{\lambda_p}, E_{\lambda_{p+1}}$ forment une somme directe. $\square$

5.2.9. Définition. — Le *polynôme caractéristique* d'une matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{K})$ est le polynôme de $\mathbb{K}[x]$ défini par

$$p_{\mathbf{A}} = \det(\mathbf{A} - x\mathbf{1}_n).$$

5.2.10. Exemples. — Le polynôme caractéristique de la matrice $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ est

$$p_{\mathbf{A}} = \begin{vmatrix} -x & 1 \\ 1 & -x \end{vmatrix} = x^2 - 1 = (x - 1)(x + 1).$$

Le polynôme caractéristique de la matrice $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ est

$$p_{\mathbf{A}} = \begin{vmatrix} 1-x & 1 \\ 1 & 1-x \end{vmatrix} = (1-x)^2 - 1 = -x(2-x).$$

5.2.11 Proposition. — Deux matrices semblables ont le même polynôme caractéristique.

Preuve. Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices semblables. Il existe donc une matrice inversible $\mathbf{P}$ telle que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Montrons que $p_{\mathbf{A}} = p_{\mathbf{B}}$. Il découle de la propriété du déterminant que :

$$\begin{aligned} p_{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}} &= \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - x\mathbf{1}_n) \\ &= \det(\mathbf{P}^{-1}(\mathbf{A} - x\mathbf{1}_n)\mathbf{P}) \\ &= (\det \mathbf{P})^{-1} \det(\mathbf{A} - x\mathbf{1}_n) \det \mathbf{P}. \end{aligned}$$

Ainsi $p_{\mathbf{A}} = p_{\mathbf{B}}$. $\square$

Par définition, un scalaire λ est valeur propre d'une matrice $\mathbf{A}$ si le déterminant de la matrice $\mathbf{A} - \lambda\mathbf{1}_n$ est nul. On a donc

5.2.12 Proposition. — Un scalaire λ est valeur propre d'une matrice $\mathbf{A}$ si, et seulement si, λ est racine de son polynôme caractéristique $p_{\mathbf{A}}$.

5.2.13. Exemple. — Le polynôme caractéristique d'une matrice de $\mathcal{M}_n(\mathbb{K})$ peut ne pas avoir de racines dans le corps $\mathbb{K}$. Par exemple, le polynôme caractéristique de la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

est $p_{\mathbf{A}} = x^2 + 1$, qui n'a pas de racines réelles, par suite $\text{Sp}_{\mathbb{R}}(\mathbf{A}) = \emptyset$. Dans le chapitre suivant, nous verrons qu'alors la matrice $\mathbf{A}$ n'est pas diagonalisable dans $\mathcal{M}_2(\mathbb{R})$, mais est diagonalisable dans $\mathcal{M}_2(\mathbb{C})$. En effet, on a $\text{Sp}_{\mathbb{C}}(\mathbf{A}) = \{-i, i\}$ et les sous-espaces propres de $\mathbf{A}$ dans $\mathbb{C}^2$ sont

$$E_i = \text{Vect}\left(\begin{bmatrix} i \\ 1 \end{bmatrix}\right) \quad \text{et} \quad E_{-i} = \text{Vect}\left(\begin{bmatrix} -i \\ 1 \end{bmatrix}\right).$$

§ 3 Trigonalisation des matrices

5.3.1. Définition. — Une matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{K})$ est dite *trigonalisable* dans $\mathcal{M}_n(\mathbb{K})$, si $\mathbf{A}$ est semblable à une matrice triangulaire supérieure de $\mathcal{M}_n(\mathbb{K})$. C'est-à-dire, s'il existe une matrice inversible $\mathbf{P}$ de $\mathcal{M}_n(\mathbb{K})$ et une matrice triangulaire supérieure $\mathbf{T}$ à coefficients dans $\mathbb{K}$ telles que

$$\mathbf{A} = \mathbf{P}\mathbf{T}\mathbf{P}^{-1}. \quad (5.2)$$

5.3.2 Théorème (Théorème de trigonalisation). — Une matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{K})$ est trigonalisable dans $\mathcal{M}_n(\mathbb{K})$ si, et seulement si, son polynôme caractéristique $p_{\mathbf{A}}$ est scindé sur $\mathbb{K}$.

Preuve. La condition est nécessaire. Si $\mathbf{A}$ est une matrice trigonalisable, par définition, elle est semblable à une matrice triangulaire supérieure :

$$\mathbf{T} = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Le polynôme caractéristique de la matrice $\mathbf{T}$ est scindé :

$$p_{\mathbf{T}} = (-1)^n (x - \lambda_1) \cdots (x - \lambda_n).$$

D'après la proposition 5.2.11, deux matrices semblables ont même polynôme caractéristique. Ainsi, $p_{\mathbf{A}} = p_{\mathbf{T}}$ et par suite le polynôme caractéristique de $\mathbf{A}$ est scindé sur $\mathbb{K}$.

La condition est suffisante. On procède par récurrence sur n . Toute matrice de $\mathcal{M}_1(\mathbb{K})$ est trigonalisable. On suppose que toute matrice de $\mathcal{M}_{n-1}(\mathbb{K})$, dont le polynôme caractéristique est scindé, est trigonalisable, montrons que cela est vrai pour toute matrice de $\mathcal{M}_n(\mathbb{K})$.

Soit $\mathbf{A} \in \mathcal{M}_n(\mathbb{K})$, telle que le polynôme $p_{\mathbf{A}}$ soit scindé sur $\mathbb{K}$. Le polynôme $p_{\mathbf{A}}$ admet donc au moins une racine λ dans $\mathbb{K}$. Considérons un vecteur propre $\mathbf{e}$ dans $\mathbb{K}^n$ associé à la valeur propre λ . Complétons le vecteur $\mathbf{e}$ en une base $\mathcal{B} = (\mathbf{e}, \mathbf{e}_2, \dots, \mathbf{e}_n)$ de $\mathbb{K}^n$. Soit $u_{\mathbf{A}}$

l'endomorphisme de $\mathbb{K}^n$ associé à la matrice $\mathbf{A}$, *i.e.*, l'endomorphisme défini, pour tout vecteur $\mathbf{x}$ de $\mathbb{K}^n$, par $u_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x}$. On a

$$u_{\mathbf{A}}(\mathbf{e}) = \mathbf{A}\mathbf{e} = \lambda\mathbf{e},$$

par suite, la matrice de l'endomorphisme $u_{\mathbf{A}}$ exprimé dans la base $\mathcal{B}$ est

$$[u_{\mathbf{A}}]_{\mathcal{B}} = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & \mathbf{B} & \\ 0 & & & \end{bmatrix},$$

où $\mathbf{B}$ est une matrice de $\mathcal{M}_{n-1}(\mathbb{K})$. La matrice $\mathbf{A}$ étant semblable à la matrice $[u_{\mathbf{A}}]_{\mathcal{B}}$, il existe une matrice inversible $\mathbf{P}$ de $\mathcal{M}_n(\mathbb{C})$, telle que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & \mathbf{B} & & \\ 0 & & & \end{bmatrix}.$$

De plus le polynôme caractéristique du bloc $\mathbf{B}$ divise le polynôme caractéristique de la matrice $\mathbf{A}$, il est donc scindé comme ce dernier. Par hypothèse de récurrence, la matrice $\mathbf{B}$ est semblable à une matrice triangulaire supérieure, il existe une matrice inversible $\mathbf{Q}$ dans $\mathcal{M}_{n-1}(\mathbb{K})$, telle que $\mathbf{t}' = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$ soit triangulaire supérieure. En multipliant par blocs, on a :

$$\begin{aligned} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{Q} & \\ 0 & & & \end{bmatrix}^{-1} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{Q} & \\ 0 & & & \end{bmatrix} &= \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q} & \\ 0 & & & \end{bmatrix} \\ &= \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & \mathbf{T}' & \\ 0 & & & \end{bmatrix}. \end{aligned}$$

En posant

$$\mathbf{R} = \mathbf{P} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{Q} & \\ 0 & & & \end{bmatrix},$$

la dernière égalité s'écrit

$$\mathbf{R}^{-1}\mathbf{A}\mathbf{R} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{Q} & \\ 0 & & & \end{bmatrix}.$$

Ainsi, $\mathbf{A}$ est semblable à une triangulaire supérieure. $\square$

D'après le théorème de D'Alembert-Gauss, tout polynôme non nul de $\mathbb{C}[x]$ est scindé sur $\mathbb{C}$. Par suite, on a

5.3.3 Proposition. — Toute matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable dans $\mathcal{M}_n(\mathbb{C})$.

Notons que toute matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{R})$ peut toujours se trigonaliser dans $\mathcal{M}_n(\mathbb{C})$. En effet, si le polynôme caractéristique de $\mathbf{A}$ est scindé sur $\mathbb{R}$, $\mathbf{A}$ est trigonalisable dans $\mathcal{M}_n(\mathbb{R})$. Sinon, le polynôme $p_{\mathbf{A}}$ est toujours scindé dans $\mathcal{M}_n(\mathbb{C})$. Il existe alors une matrice inversible $\mathbf{P}$ et une matrice triangulaire $\mathbf{T}$ de $\mathcal{M}_n(\mathbb{C})$ telles que $\mathbf{A} = \mathbf{P}\mathbf{T}\mathbf{P}^{-1}$.

5.3.4. Exemple. — La matrice suivante de $\mathcal{M}_4(\mathbb{R})$

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

admet pour polynôme caractéristique

$$p_{\mathbf{A}} = (x^2 + 1)^2.$$

Ce polynôme n'est pas scindé dans $\mathbb{R}[x]$, la matrice $\mathbf{A}$ n'est donc pas trigonalisable dans $\mathcal{M}_4(\mathbb{R})$. Cependant, il est scindé dans $\mathbb{C}[x]$:

$$p_{\mathbf{A}} = (x - i)^2(x + i)^2.$$

La matrice est trigonalisable. Posons

$$\mathbf{P} = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -i & 0 & i & i \\ 0 & 1 & 0 & 1 \\ 0 & -i & 0 & i \end{bmatrix}.$$

Le premier et troisième vecteur colonne de la matrice $\mathbf{P}$ sont des vecteurs propres associés aux valeurs propres i et $-i$ respectivement. Les deux autres vecteurs colonnes complètent ces vecteurs en une base de trigonalisation. On a

$$\mathbf{A} = \mathbf{P} \begin{bmatrix} i & 1 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 1 \\ 0 & 0 & 0 & -i \end{bmatrix} \mathbf{P}^{-1}, \quad \text{avec} \quad \mathbf{P}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & i & 1 & 0 \\ 0 & 0 & 1 & i \\ 1 & -i & 0 & i \\ 0 & 0 & 1 & -i \end{bmatrix}.$$

5.3.5. Somme et produit des valeurs propres. — Si une matrice $\mathbf{A}$ est trigonalisable, semblable à une matrice triangulaire supérieure $\mathbf{T}$, alors les valeurs propres de $\mathbf{A}$ étant les racines du polynôme $p_{\mathbf{A}}$, sont aussi les coefficients de la diagonale de la matrice $\mathbf{T}$.

Étant donnée une matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{C})$, alors son polynôme caractéristique est scindé sur $\mathbb{C}$:

$$p_{\mathbf{A}} = (-1)^n(x - \lambda_1) \dots (x - \lambda_n).$$

La matrice $\mathbf{A}$ est semblable à une matrice triangulaire $\mathbf{T}$, *i.e.*, il existe une matrice inversible $\mathbf{P}$ telle que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Étant semblables, les matrices $\mathbf{A}$ et $\mathbf{T}$ ont même trace et même déterminant, on en déduit que la trace (resp. le déterminant) de $\mathbf{A}$ est égale à la somme (resp. le produit) des valeurs propres, comptées avec leur ordre de multiplicité. Précisément, on a

5.3.6 Proposition. — Soit $\mathbf{A}$ une matrice de $\mathcal{M}_n(\mathbb{C})$ de polynôme caractéristique

$$p_{\mathbf{A}} = (-1)^n (x - \lambda_1)^{n_1} \dots (x - \lambda_p)^{n_p},$$

où n_i désigne l'ordre de multiplicité de la valeur propre λ_i dans le polynôme caractéristique. Alors,

i) $\text{trace}(\mathbf{A}) = n_1 \lambda_1 + \dots + n_p \lambda_p,$

ii) $\det(\mathbf{A}) = \lambda_1^{n_1} \dots \lambda_p^{n_p}.$

5.3.7. Exemples. — Dans l'exemple 5.2.13, on a montré que la matrice $\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, possède deux valeurs propres $-i$ et i ; la somme de ces valeurs propres est égale à la trace de $\mathbf{A}$ et leur produit est le déterminant de $\mathbf{A}$.

§ 4 Diagonalisation des matrices

5.4.1. Matrices diagonalisables. — Une matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{K})$ est dite *diagonalisable* dans $\mathcal{M}_n(\mathbb{K})$, si elle est semblable à une matrice diagonale de $\mathcal{M}_n(\mathbb{K})$. C'est-à-dire, s'il existe une matrice inversible $\mathbf{P}$ de $\mathcal{M}_n(\mathbb{K})$ et une matrice diagonale $\mathbf{D}$ à coefficients dans $\mathbb{K}$ telles que

$$\mathbf{A} = \mathbf{P} \mathbf{D} \mathbf{P}^{-1}. \quad (5.3)$$

Les matrices $\mathbf{A}$ et $\mathbf{D}$ de la décomposition (5.3) étant semblables, d'après la proposition 5.2.11, elles ont le même polynôme caractéristique. Il s'ensuit que la diagonale de la matrice $\mathbf{D}$ est formée des valeurs propres de $\mathbf{A}$.

5.4.2 Exercice. — Montrer que la matrice $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ est diagonalisable dans $\mathcal{M}_2(\mathbb{R})$. On a

$$\mathbf{A} = \mathbf{P} \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \mathbf{P}^{-1},$$

avec

$$\mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad \text{et} \quad \mathbf{P}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

5.4.3 Proposition. — Une matrice $\mathbf{A}$ de $\mathcal{M}_n(\mathbb{K})$ est diagonalisable si, et seulement si, il existe une base de $\mathbb{K}^n$ formée de vecteurs propres de $\mathbf{A}$.

Preuve. Supposons qu'il existe une base $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ de $\mathbb{K}^n$ composée de vecteurs propres de $\mathbf{A}$. Considérons la matrice $\mathbf{P}$ dont les colonnes sont formées par les éléments de cette base :

$$\mathbf{P} = \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix}.$$

Les vecteurs $\mathbf{x}_1, \dots, \mathbf{x}_n$ formant une base de $\mathbb{K}^n$, la matrice $\mathbf{P}$ est inversible et on a

$$\begin{aligned} \mathbf{AP} &= \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ \mathbf{Ax}_1 & \mathbf{Ax}_2 & \cdots & \mathbf{Ax}_n \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \\ &= \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ \lambda_1 \mathbf{x}_1 & \lambda_2 \mathbf{x}_2 & \cdots & \lambda_n \mathbf{x}_n \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \\ &= \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \end{aligned}$$

Ainsi

$$\mathbf{AP} = \mathbf{P} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}. \quad (5.4)$$

Par suite, la matrice $\mathbf{A}$ est diagonalisable.

Inversement, si $\mathbf{A}$ est diagonalisable, il existe une matrice inversible $\mathbf{P}$ telle que $\mathbf{P}$ satisfait l'égalité (5.4). Pour les mêmes raisons, les vecteurs colonnes de $\mathbf{P}$ forment une base de vecteurs propres de $\mathbb{K}^n$. $\square$

5.4.4 Proposition (Condition suffisante de diagonalisation). — Soit $\mathbf{A}$ une matrice de $\mathcal{M}_n(\mathbb{K})$. Si le polynôme caractéristique de $\mathbf{A}$ est scindé sur $\mathbb{K}$ et possède toutes ses racines simples, alors $\mathbf{A}$ est diagonalisable dans $\mathcal{M}_n(\mathbb{K})$.

Preuve. Supposons que le polynôme caractéristique $p_{\mathbf{A}}$ soit scindé à racines simples :

$$p_{\mathbf{A}} = (-1)^n (x - \lambda_1) \cdots (x - \lambda_n),$$

avec $\lambda_i \neq \lambda_j$, si $i \neq j$. Pour tout $i \in \llbracket 1, n \rrbracket$, λ_i étant valeur propre de $\mathbf{A}$, il existe un vecteur propre $\mathbf{x}_i$ de $\mathbf{A}$ associé. D'après la proposition 5.2.8, les sous-espaces propres de $\mathbf{A}$ formant une somme directe, la famille $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ est libre. Elle possède n éléments, c'est donc une base de $\mathbb{K}^n$. $\square$

On en déduit le résultat suivant :

5.4.5 Corollaire. — Toute matrice de $\mathcal{M}_n(\mathbb{K})$ qui admet n valeurs propres distinctes est diagonalisable.

§ 5 Diagonalisation et applications

5.5.1. Calcul des puissances d'une matrice diagonalisable. — Une première application classique de la diagonalisation est le calcul des puissances d'une matrice. Soit $\mathbf{A}$ une matrice diagonalisable de $\mathcal{M}_n(\mathbb{K})$. Il existe une matrice diagonale

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix}$$

et une matrice inversible $\mathbf{P}$ telles que : $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Alors $\mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}$, pour tout entier naturel k , d'où

$$\mathbf{A}^k = \mathbf{P} \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_p^k \end{bmatrix} \mathbf{P}^{-1}$$

5.5.2. Exemple. — Les puissances successives de la matrice $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$ sont

$$\mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} = \begin{bmatrix} 2^{k+1} - 3^k & 2^{k+1} - 2 \cdot 3^k \\ -2^k + 3^k & -2^k + 2 \cdot 3^k \end{bmatrix}$$

avec

$$\mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}, \quad \mathbf{P}^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}.$$

5.5.3. Résolution des systèmes différentiels linéaires. — Une autre application classique de la diagonalisation d'une matrice est la résolution des systèmes différentiels linéaires. On se propose de résoudre les systèmes différentiels linéaires de la forme :

$$\begin{cases} \frac{dx_1(t)}{dt} = a_1^1 x_1(t) + \cdots + a_1^n x_n(t) \\ \vdots \\ \frac{dx_n(t)}{dt} = a_n^1 x_1(t) + \cdots + a_n^n x_n(t) \end{cases}$$

où les a_i^j sont des réels et les x_i des fonctions réelles à valeurs réelles. Un tel système différentiel prend la forme matricielle suivante :

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t), \tag{5.5}$$

avec

$$\mathbf{A} = \begin{bmatrix} a_1^1 & \cdots & a_1^n \\ \vdots & & \vdots \\ a_n^1 & \cdots & a_n^n \end{bmatrix}, \quad \mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

et où $\frac{d\mathbf{x}(t)}{dt}$ désigne la dérivée du vecteur $\mathbf{x}(t)$.

Supposons que la matrice $\mathbf{A}$ soit diagonalisable (nous aborderons le cas des systèmes différentiels avec $\mathbf{A}$ non diagonalisable plus tard dans le cours), il existe une matrice $\mathbf{D}$ diagonale et $\mathbf{P}$ inversible telles que

$$\mathbf{D} = \mathbf{P}^{-1} \mathbf{A} \mathbf{P}.$$

La matrice $\mathbf{P}$ désigne un changement de base de $\mathbb{R}^n$: si $\mathbf{x}(t)$ est le vecteur colonne exprimant un vecteur $\mathbf{x}(t)$ dans la base initiale et $\mathbf{y}(t)$ celui exprimant $\mathbf{x}(t)$ dans la nouvelle base. On fait le changement de variable $\mathbf{y}(t) = \mathbf{P}^{-1} \mathbf{x}(t)$. D'où

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{P}^{-1} \frac{d\mathbf{x}(t)}{dt},$$

donc

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{P}^{-1} \mathbf{A} \mathbf{x}(t) = \mathbf{P}^{-1} \mathbf{A} \mathbf{P} \mathbf{y}(t) = \mathbf{D} \mathbf{y}(t).$$

Le système est donc équivalent au système :

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{D} \mathbf{y}(t),$$

qui est facile à intégrer, car $\mathbf{D}$ est diagonale. En effet, si

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix}$$

les solutions de cette équation sont les vecteurs $\mathbf{y}(t)$ dont la i -ième composante est

$$y_i(t) = e^{\lambda_i t} y_i(0).$$

Il suffit alors de calculer $\mathbf{x}(t)$ en calculant $\mathbf{x}(t) = \mathbf{P} \mathbf{y}(t)$.

5.5.4. Exemple. — Soit à résoudre le système différentiel suivant :

$$\begin{cases} \frac{dx(t)}{dt} = x(t) - y(t), \\ \frac{dy(t)}{dt} = 2x(t) + 4y(t) \end{cases} \quad (5.6)$$

où $x, y : \mathbb{R} \longrightarrow \mathbb{R}$ sont deux fonctions réelles. On pose

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{et} \quad \mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}.$$

Le système $\frac{d\mathbf{y}(t)}{dt} = \mathbf{D} \mathbf{y}(t)$, avec $\mathbf{y}(t) = \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}$ s'écrit sous la forme :

$$\begin{cases} \frac{du(t)}{dt} = 2u(t) \\ \frac{dv(t)}{dt} = 3v(t) \end{cases}$$

Ces deux équations ont pour solution

$$\begin{cases} u(t) = \beta_1 e^{2t} \\ v(t) = \beta_2 e^{3t} \end{cases}$$

où β_1 et β_2 sont deux constantes réelles. On en déduit que le système (5.6) admet pour solution

$$\begin{cases} x(t) = \beta_1 e^{2t} + \beta_2 e^{3t} \\ y(t) = -\beta_1 e^{2t} - 2\beta_2 e^{3t} \end{cases}$$